

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2003, 13:24:45 ; Search time 31.3571 Seconds  
(without alignments)  
46.744 Million cell updates/sec

Title: US-09-909-164-5  
Perfect score: 52  
Sequence: 1 EGVVPXGMSYS 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues  
Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 75 summaries

Database : A\_Geneseq\_101002.\*  
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Query Match	Length	ID	Description
1	50	96.2	11	23	ABB80521	Hepatitis C virus
2	50	96.2	11	23	ABB80522	Hepatitis C virus
3	50	96.2	11	23	ABB80525	Hepatitis C virus
4	50	96.2	11	23	ABB80526	Hepatitis C virus
5	50	96.2	11	23	ABB80559	Hepatitis C virus
6	50	96.2	11	23	ABB80563	Hepatitis C virus
7	50	96.2	11	23	ABB80564	Hepatitis C virus
8	50	96.2	11	23	ABB80565	Hepatitis C virus
9	50	96.2	11	23	ABB80566	Hepatitis C virus
10	50	96.2	11	23	ABB80567	Hepatitis C virus

10/8/02  
(after P10)  
1st entry of  
ALL APPLICATIONS SEQ'S  
ONLY APPLICATION w/ Close Homology  
100%

11	50	96.2	11	23	ABB80568	Hepatitis C virus
12	46	88.5	11	23	ABB80524	Hepatitis C virus
13	46	88.5	11	23	ABB80528	Hepatitis C virus
14	46	88.5	11	23	ABB80529	Hepatitis C virus
15	46	88.5	11	23	ABB80561	Hepatitis C virus
16	46	88.5	11	23	ABB80562	Hepatitis C virus
17	45	86.5	11	23	ABB80523	Hepatitis C virus
18	45	86.5	11	23	ABB80527	Hepatitis C virus
19	45	86.5	11	23	ABB80535	Hepatitis C virus
20	45	86.5	11	23	ABB80536	Hepatitis C virus
21	45	86.5	11	23	ABB80539	Hepatitis C virus
22	45	86.5	11	23	ABB80540	Hepatitis C virus
23	45	86.5	11	23	ABB80558	Hepatitis C virus
24	45	86.5	11	23	ABB80560	Hepatitis C virus
25	44	84.6	11	23	ABB80544	Hepatitis C virus
26	44	84.6	11	23	ABB80545	Hepatitis C virus
27	44	84.6	11	23	ABB80549	Hepatitis C virus
28	44	84.6	11	23	ABB80552	Hepatitis C virus
29	44	84.6	11	23	ABB80553	Hepatitis C virus
30	42	80.8	11	23	ABB80530	Hepatitis C virus
31	41	78.8	11	23	ABB80538	Hepatitis C virus
32	41	78.8	11	23	ABB80542	Hepatitis C virus
33	41	78.8	11	23	ABB80543	Hepatitis C virus
34	40	76.9	11	23	ABB80537	Hepatitis C virus
35	40	76.9	11	23	ABB80541	Hepatitis C virus
36	40	76.9	11	23	ABB80547	Hepatitis C virus
37	40	76.9	11	23	ABB80548	Hepatitis C virus
38	40	76.9	11	23	ABB80551	Hepatitis C virus
39	40	76.9	11	23	ABB80556	Hepatitis C virus
40	40	76.9	11	23	ABB80557	Hepatitis C virus
41	40	76.9	20	20	AAU76810	Novel human diapo
42	40	76.9	1022	22	ABG03621	Novel human diapo
43	40	76.9	1022	22	ABG05826	Novel human diapo
44	40	76.9	1022	22	ABG08173	Novel human diapo
45	39	75.0	11	23	ABB80546	Hepatitis C virus
46	39	75.0	11	23	ABB80550	Hepatitis C virus
47	39	75.0	11	23	ABB80554	Hepatitis C virus
48	39	75.0	11	23	ABB80555	Hepatitis C virus
49	38	73.1	11	23	ABB80533	Hepatitis C virus
50	38	73.1	11	23	ABB80534	Hepatitis C virus
51	38	73.1	3472	21	AA190913	Cenarchaeum symbio
52	37	71.2	11	23	ABB80531	Hepatitis C virus
53	37	71.2	11	23	ABB80532	Hepatitis C virus
54	36	69.2	244	21	AA12881	Murine JNK3 bindin
55	36	69.2	484	21	AA12882	Murine JNK3 bindin
56	35	67.3	11	18	AAW99288	Peptide N424 from
57	34	65.4	842	21	AA144359	P. chrysogenum sut
58	34	65.4	947	21	AA14378	Pinus radiata cell
59	34	65.4	1070	22	AA125105	Human novel protei
60	33	63.5	12	21	AA183772	HCV NS3 protease s
61	33	63.5	12	21	AA183774	HCV NS3 protease s
62	33	63.5	13	18	AAW99276	Peptide D4 from WO
63	33	63.5	13	18	AAW33286	Peptide 5 used in
64	33	63.5	14	18	AAW99275	Peptide D3 from WO
65	33	63.5	14	18	AAW99277	Peptide C0 from WO
66	33	63.5	14	18	AAW33285	Peptide 4 used in
67	33	63.5	14	18	AAW33287	Peptide 6 used in
68	33	63.5	16	18	AAW99274	Peptide D2 from WO
69	33	63.5	16	18	AAW33284	Peptide 3 used in
70	33	63.5	18	18	AAW99273	Peptide D1 from WO
71	33	63.5	18	18	AAW33283	Peptide 2 used in
72	33	63.5	18	19	AAW71281	Cleavable substrate
73	33	63.5	20	18	AAW99272	Peptide PS from WO
74	33	63.5	20	18	AAW33282	Peptide 1 used in
75	33	63.5	20	20	AA14511	HCV NS3 protease t

ALIGNMENTS

RESULT 1  
ABB80521  
ID ABB80521 standard; peptide: 11 AA.

XX AC ABB80521-  
 XX DT 08-OCT-2002 (first entry)  
 XX DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #1.  
 XX KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
 XX KW virucide.  
 XX OS Synthetic.  
 XX FH Key Location/Qualifiers  
 XX FT Modified-site 1 /note= "N-terminal acetyl"  
 XX FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with  
 XX FT residue 7"  
 XX FT Modified-site 11 /note= "C-terminal amide"  
 XX PN WO200208251-A2.  
 XX PD 31-JAN-2002.  
 XX PF 19-JUL-2001; 2001WO-US23169.  
 XX PR 21-JUL-2000; 2000US-220101P.  
 XX PA (CORV-) CORVAS INT INC.  
 XX PI Lim-wilby M, Levy OE, Brunck TK;  
 XX DR WPI; 2002-361643/39.  
 XX CC Novel peptide compound having hepatitis C virus protease inhibitory  
 XX FT activity useful for treating disorders associated with hepatitis C  
 XX FT virus protease  
 XX PS Claim 17; Page 64; 69pp; English.  
 XX CC The sequence represents a peptide compound of the invention having  
 XX CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
 XX CC invention are alpha-ketoamide peptide analogues. The peptides have  
 XX CC virucide activity, and are useful for treating and in the manufacture of  
 XX CC a medicament to treat disorders associated with HCV protease. A  
 XX CC pharmaceutical composition comprising the peptide as an active ingredient  
 XX CC is useful for treating disorders associated with hepatitis C virus.  
 XX SQ Sequence 11 AA;  
 XX Query Match 96.2%; Score 50; DB 23; Length 11;  
 XX Best Local Similarity 100.0%; Pred. No. 0.0011;  
 XX Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 EEVVPXGMSYS 11  
 Db 1 EEVVPXGMSYS 11  
 RESULT 2  
 ABB80522  
 ID ABB80522 standard; peptide; 11 AA.  
 XX AC ABB80522;  
 XX DT 08-OCT-2002 (first entry)  
 XX DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #2.  
 XX KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
 XX KW virucide.

OS Synthetic.  
 XX Key Location/Qualifiers  
 XX FT Modified-site 1 /note= "N-terminal acetyl"  
 XX FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with  
 XX FT residue 7"  
 XX FT Misc-difference 9 /note= "D-form residue"  
 XX FT Modified-site 11 /note= "C-terminal amide"  
 XX PN WO200208251-A2.  
 XX PD 31-JAN-2002.  
 XX PF 19-JUL-2001; 2001WO-US23169.  
 XX PR 21-JUL-2000; 2000US-220101P.  
 XX PA (CORV-) CORVAS INT INC.  
 XX PI Lim-wilby M, Levy OE, Brunck TK;  
 XX DR WPI; 2002-361643/39.  
 XX CC Novel peptide compound having hepatitis C virus protease inhibitory  
 XX FT activity useful for treating disorders associated with hepatitis C  
 XX FT virus protease  
 XX PS Claim 17; Page 64; 69pp; English.  
 XX CC The sequence represents a peptide compound of the invention having  
 XX CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
 XX CC invention are alpha-ketoamide peptide analogues. The peptides have  
 XX CC virucide activity, and are useful for treating and in the manufacture of  
 XX CC a medicament to treat disorders associated with HCV protease. A  
 XX CC pharmaceutical composition comprising the peptide as an active ingredient  
 XX CC is useful for treating disorders associated with hepatitis C virus.  
 XX SQ Sequence 11 AA;  
 XX Query Match 96.2%; Score 50; DB 23; Length 11;  
 XX Best Local Similarity 100.0%; Pred. No. 0.0011;  
 XX Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 EEVVPXGMSYS 11  
 Db 1 EEVVPXGMSYS 11  
 RESULT 3  
 ABB80525  
 ID ABB80525 standard; peptide; 11 AA.  
 XX AC ABB80525;  
 XX DT 08-OCT-2002 (first entry)  
 XX DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #5.  
 XX KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
 XX KW virucide.  
 XX OS Synthetic.  
 XX FH Key Location/Qualifiers  
 XX FT Modified-site 1 /note= "N-terminal acetyl"  
 XX FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with  
 XX FT residue 7"

FT. Misc-difference 8 /note= "D-form residue"  
FT Modified-site 11  
FT Modified-site /note= "C-terminal amide"

WO200208251-A2.

31-JAN-2002.

19-JUL-2001; 2001WO-US23169.

21-JUL-2000; 2000US-220101P.

(CORV-) CORVAS INT INC.

Lim-wilby M, Levy OE, Brunck TK;

WPI; 2002-361643/39.

Novel peptide compound having hepatitis C virus protease inhibitory activity useful for treating disorders associated with hepatitis C virus protease

Claim 17; Page 64; 69pp; English.

The sequence represents a peptide compound of the invention having hepatitis C virus (HCV) protease inhibitory activity. The peptides of the invention are alpha-ketoamide peptide analogues. The peptides have virucide activity, and are useful for treating and in the manufacture of a medicament to treat disorders associated with HCV protease. A pharmaceutical composition comprising the peptide as an active ingredient is useful for treating disorders associated with hepatitis C virus.

Sequence 11 AA;

Query Match 96.2%; Score 50; DB 23; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.0011;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEVVPXGMSYS 11  
Db 1 EEVVPXGMSYS 11

RESULT 4  
ABB80526  
ID ABB80526 standard; peptide; 11 AA.

AC ABB80526;

DT 08-OCT-2002 (first entry)

Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #6.

Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide; virucide.

Synthetic.

Key Location/Qualifiers

Modified-site 1 /note= "N-terminal acetyl"

Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with residue 7"

Misc-difference 8

Misc-difference 9 /note= "D-form residue"

Misc-difference 11 /note= "D-form residue"

Modified-site 11 /note= "C-terminal amide"

WO200208251-A2.

XX 31-JAN-2002.

19-JUL-2001; 2001WO-US23169.

21-JUL-2000; 2000US-220101P.

(CORV-) CORVAS INT INC.

Lim-wilby M, Levy OE, Brunck TK;

WPI; 2002-361643/39.

Novel peptide compound having hepatitis C virus protease inhibitory activity useful for treating disorders associated with hepatitis C virus protease

Claim 17; Page 64; 69pp; English.

The sequence represents a peptide compound of the invention having hepatitis C virus (HCV) protease inhibitory activity. The peptides of the invention are alpha-ketoamide peptide analogues. The peptides have virucide activity, and are useful for treating and in the manufacture of a medicament to treat disorders associated with HCV protease. A pharmaceutical composition comprising the peptide as an active ingredient is useful for treating disorders associated with hepatitis C virus.

Sequence 11 AA;

Query Match 96.2%; Score 50; DB 23; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.0011;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEVVPXGMSYS 11  
Db 1 EEVVPXGMSYS 11

RESULT 5  
ABB80559  
ID ABB80559 standard; peptide; 11 AA.

AC ABB80559;

DT 08-OCT-2002 (first entry)

Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #39.

Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide; virucide.

Synthetic.

Key Location/Qualifiers

Modified-site 1 /note= "N-terminal acetyl"

Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with residue 7"

Misc-difference 8 /note= "D-form residue"

Modified-site 8 /note= "Oxymethionine"

Modified-site 11 /note= "C-terminal amide"

WO200208251-A2.

31-JAN-2002.

19-JUL-2001; 2001WO-US23169.

21-JUL-2000; 2000US-220101P.

XX PA (CORV-) CORVAS INT INC.  
 XX PI Lim-wilby M, Levy OE, Brunck TK;  
 XX DR WPI; 2002-361643/39.  
 XX PT Novel peptide compound having hepatitis C virus protease inhibitory  
 PT activity useful for treating disorders associated with hepatitis C  
 PT virus protease -  
 XX PS  
 XX PS Claim 17; Page 65; 69pp; English.  
 XX CC The sequence represents a peptide compound of the invention having  
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
 CC invention are alpha-ketoamide peptide analogues. The peptides have  
 CC virucide activity, and are useful for treating and in the manufacture of  
 CC a medicament to treat disorders associated with HCV protease. A  
 CC pharmaceutical composition comprising the peptide as an active ingredient  
 CC is useful for treating disorders associated with hepatitis C virus.  
 XX Sequence 11 AA;

Query Match 96.2%; Score 50; DB 23; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 0.0011;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 EEVVPXGMSYS 11  
 |||||  
 Db 1 EEVVPXGMSYS 11

RESULT 6  
 ABB80563  
 ID ABB80563 standard; peptide; 11 AA.  
 AC ABB80563;  
 XX 08-OCT-2002 (first entry)  
 DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #43.  
 XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
 KW virucide.  
 KW Synthetic.  
 OS

Key Location/Qualifiers  
 FH Modified-site 1 /note= "N-terminal acetyl"  
 FT Modified-site 6 /note= "Valyl carbonyl forming keto-amide linkage with  
 FT residue 7"  
 FT Modified-site 11 /note= "C-terminal amide"  
 FT

WO200208251-A2.  
 31-JAN-2002.  
 19-JUL-2001; 2001WO-US23169.  
 21-JUL-2000; 2000US-220101P.  
 (CORV-) CORVAS INT INC.  
 Lim-wilby M, Levy OE, Brunck TK;  
 WPI; 2002-361643/39.

Novel peptide compound having hepatitis C virus protease inhibitory  
 activity useful for treating disorders associated with hepatitis C  
 virus protease -  
 PT

XX PS Claim 17; Page 65; 69pp; English.  
 XX CC The sequence represents a peptide compound of the invention having  
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
 CC invention are alpha-ketoamide peptide analogues. The peptides have  
 CC virucide activity, and are useful for treating and in the manufacture of  
 CC a medicament to treat disorders associated with HCV protease. A  
 CC pharmaceutical composition comprising the peptide as an active ingredient  
 CC is useful for treating disorders associated with hepatitis C virus.  
 XX Sequence 11 AA;

Query Match 96.2%; Score 50; DB 23; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 0.0011;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 EEVVPXGMSYS 11  
 |||||  
 Db 1 EEVVPXGMSYS 11

RESULT 7  
 ABB80564  
 ID ABB80564 standard; peptide; 11 AA.  
 XX ABB80564;  
 AC ABB80564;  
 XX 08-OCT-2002 (first entry)  
 DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #44.  
 XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
 KW virucide.  
 KW Synthetic.  
 OS

Key Location/Qualifiers  
 FH Modified-site 1 /note= "N-terminal acetyl"  
 FT Modified-site 6 /note= "Leucyl carbonyl forming keto-amide linkage with  
 FT residue 7"  
 FT Modified-site 11 /note= "C-terminal amide"  
 FT

WO200208251-A2.  
 31-JAN-2002.  
 19-JUL-2001; 2001WO-US23169.  
 21-JUL-2000; 2000US-220101P.  
 (CORV-) CORVAS INT INC.  
 Lim-wilby M, Levy OE, Brunck TK;  
 WPI; 2002-361643/39.  
 Novel peptide compound having hepatitis C virus protease inhibitory  
 activity useful for treating disorders associated with hepatitis C  
 virus protease -

Claim 17; Page 65; 69pp; English.  
 The sequence represents a peptide compound of the invention having  
 hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
 invention are alpha-ketoamide peptide analogues. The peptides have  
 virucide activity, and are useful for treating and in the manufacture of  
 a medicament to treat disorders associated with HCV protease. A  
 pharmaceutical composition comprising the peptide as an active ingredient  
 is useful for treating disorders associated with hepatitis C virus.



```

XX SQ Sequence 11 AA;
Query Match 96.2%; Score 50; DB 23; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.0011;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEVVPXGMSYS 11
   |||||
Db 1 EEVVPXGMSYS 11

RESULT 8
ABB80565
ID ABB80565 standard; peptide; 11 AA.
XX
XX ABB80565;
XX
DT 08-OCT-2002 (first entry)
XX
DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #45.
XX
KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
KW virucide.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT /note= "N-terminal acetyl"
FT Modified-site 6
FT /note= "Norleucyl carbonyl forming keto-amide linkage
FT with residue 7"
FT Modified-site 11
FT /note= "C-terminal amide"
XX
XX WO200208251-A2.
XX
XX 31-JAN-2002.
XX
PF 19-JUL-2001; 2001WO-US23169.
XX
XX 21-JUL-2000; 2000US-220101P.
XX
XX (CORV-) CORVAS INT INC.
XX
PI Lim-wilby M, Levy OE, Brunck TK;
XX
DR WPI; 2002-361643/39.
XX
PT Novel peptide compound having hepatitis C virus protease inhibitory
PT activity useful for treating disorders associated with hepatitis C
PT virus protease
XX
XX Claim 17; Page 65; 69pp; English.
XX
CC The sequence represents a peptide compound of the invention having
CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
CC invention are alpha-ketoamide peptide analogues. The peptides have
CC virucide activity, and are useful for treating and in the manufacture of
CC a medicament to treat disorders associated with HCV protease. A
CC pharmaceutical composition comprising the peptide as an active ingredient
CC is useful for treating disorders associated with hepatitis C virus.
XX
XX SQ Sequence 11 AA;
Query Match 96.2%; Score 50; DB 23; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.0011;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEVVPXGMSYS 11
   |||||
Db 1 EEVVPXGMSYS 11

RESULT 9
ABB80566
ID ABB80566 standard; peptide; 11 AA.
XX
XX ABB80566;
XX
DT 08-OCT-2002 (first entry)
XX
DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #46.
XX
KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
KW virucide.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT /note= "N-terminal acetyl"
FT Modified-site 6
FT /note= "2-aminoisobutyryl carbonyl residue forming a
FT keto-amide linkage with residue 7"
FT Modified-site 11
FT /note= "C-terminal amide"
XX
XX WO200208251-A2.
XX
XX 31-JAN-2002.
XX
PF 19-JUL-2001; 2001WO-US23169.
XX
XX 21-JUL-2000; 2000US-220101P.
XX
XX (CORV-) CORVAS INT INC.
XX
PI Lim-wilby M, Levy OE, Brunck TK;
XX
DR WPI; 2002-361643/39.
XX
PT Novel peptide compound having hepatitis C virus protease inhibitory
PT activity useful for treating disorders associated with hepatitis C
PT virus protease
XX
XX Claim 17; Page 65; 69pp; English.
XX
CC The sequence represents a peptide compound of the invention having
CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
CC invention are alpha-ketoamide peptide analogues. The peptides have
CC virucide activity, and are useful for treating and in the manufacture of
CC a medicament to treat disorders associated with HCV protease. A
CC pharmaceutical composition comprising the peptide as an active ingredient
CC is useful for treating disorders associated with hepatitis C virus.
XX
XX SQ Sequence 11 AA;
Query Match 96.2%; Score 50; DB 23; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.0011;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEVVPXGMSYS 11
   |||||
Db 1 EEVVPXGMSYS 11

RESULT 10
ABB80567
ID ABB80567 standard; peptide; 11 AA.
XX
XX ABB80567;
XX
DT 08-OCT-2002 (first entry)
XX

```

DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #47.  
 XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
 KW virucide.  
 XX Synthetic.  
 XX OS  
 XX Key Location/Qualifiers  
 FH Modified-site 1  
 FT /note= "N-terminal acetyl"  
 FT Modified-site 6  
 FT /note= "(s,s)allothreonyl carbonyl residue forming a  
 FT keto-amide linkage with residue 7"  
 FT Modified-site 11  
 FT /note= "C-terminal amide"  
 XX WO200208251-A2.  
 XX 31-JAN-2002.  
 XX 19-JUL-2001; 2001WO-US23169.  
 XX 21-JUL-2000; 2000US-220101P.  
 XX (CORV-) CORVAS INT INC.  
 XX Lim-wilby M, Levy OE, Brunck TK;  
 XX WPI; 2002-361643/39.  
 XX Novel peptide compound having hepatitis C virus protease inhibitory  
 PT activity useful for treating disorders associated with hepatitis C  
 PT virus protease  
 XX Claim 17; Page 65; 69pp; English.  
 XX The sequence represents a peptide compound of the invention having  
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
 CC invention are alpha-ketoamide peptide analogues. The peptides have  
 CC virucide activity, and are useful for treating and in the manufacture of  
 CC a medicament to treat disorders associated with HCV protease. A  
 CC pharmaceutical composition comprising the peptide as an active ingredient  
 CC is useful for treating disorders associated with hepatitis C virus.  
 XX SQ Sequence 11 AA;  
 XX Query Match 96.2%; Score 50; DB 23; Length 11;  
 XX Best Local Similarity 100.0%; Pred. No. 0.0011;  
 XX Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 EEVVPXGMSYS 11  
 DB | | | | | | | | | | | |  
 1 EEVVPXGMSYS 11  
 RESULT 11  
 ABB80568  
 ID ABB80568 standard; peptide; 11 AA.  
 XX ABB80568;  
 XX 08-OCT-2002 (first entry)  
 XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #48.  
 XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
 KW virucide.  
 XX Synthetic.  
 XX OS  
 XX Key Location/Qualifiers  
 FH Modified-site 1  
 FT /note= "N-terminal acetyl"

FT Modified-site 6  
 FT /note= "Alpha-propionyl-glycyl-carbonyl residue forming  
 FT a keto-amide linkage with residue 7"  
 FT Modified-site 11  
 FT /note= "C-terminal amide"  
 XX WO200208251-A2.  
 XX 31-JAN-2002.  
 XX 19-JUL-2001; 2001WO-US23169.  
 XX 21-JUL-2000; 2000US-220101P.  
 XX (CORV-) CORVAS INT INC.  
 XX Lim-wilby M, Levy OE, Brunck TK;  
 XX WPI; 2002-361643/39.  
 XX Novel peptide compound having hepatitis C virus protease inhibitory  
 PT activity useful for treating disorders associated with hepatitis C  
 PT virus protease  
 XX Claim 17; Page 65; 69pp; English.  
 XX The sequence represents a peptide compound of the invention having  
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
 CC invention are alpha-ketoamide peptide analogues. The peptides have  
 CC virucide activity, and are useful for treating and in the manufacture of  
 CC a medicament to treat disorders associated with HCV protease. A  
 CC pharmaceutical composition comprising the peptide as an active ingredient  
 CC is useful for treating disorders associated with hepatitis C virus.  
 XX SQ Sequence 11 AA;  
 XX Query Match 96.2%; Score 50; DB 23; Length 11;  
 XX Best Local Similarity 100.0%; Pred. No. 0.0011;  
 XX Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 EEVVPXGMSYS 11  
 DB | | | | | | | | | | | |  
 1 EEVVPXGMSYS 11  
 RESULT 12  
 ABB80524  
 ID ABB80524 standard; peptide; 11 AA.  
 XX ABB80524;  
 XX 08-OCT-2002 (first entry)  
 XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #4.  
 XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
 KW virucide.  
 XX Synthetic.  
 XX OS  
 XX Key Location/Qualifiers  
 FH Modified-site 1  
 FT /note= "N-terminal acetyl"  
 FT Modified-site 6  
 FT /note= "Norvalyl carbonyl forming keto-amide linkage with  
 FT residue 7"  
 FT Misc-difference 9  
 FT /note= "D-form residue"  
 FT Modified-site 11  
 FT /note= "C-terminal amide"  
 XX WO200208251-A2.

PD 31-JAN-2002.  
 XX  
 PF 19-JUL-2001; 2001WO-US23169.  
 XX  
 PR 21-JUL-2000; 2000US-220101P.  
 XX  
 PA (CORV-) CORVAS INT INC.  
 XX  
 PI Lim-wilby M, Levy OE, Brunck TK;  
 XX  
 DR WPI; 2002-361643/39.  
 XX  
 PT Novel peptide compound having hepatitis C virus protease inhibitory  
 PT activity useful for treating disorders associated with hepatitis C  
 PT virus protease  
 XX  
 PS Claim 17; Page 64; 69pp; English.  
 XX  
 CC The sequence represents a peptide compound of the invention having  
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
 CC invention are alpha-ketoamide peptide analogues. The peptides have  
 CC virucide activity, and are useful for treating and in the manufacture of  
 CC a medicament to treat disorders associated with HCV protease. A  
 CC pharmaceutical composition comprising the peptide as an active ingredient  
 CC is useful for treating disorders associated with hepatitis C virus.  
 XX  
 SQ Sequence 11 AA;  
 Query Match 88.5%; Score 46; DB 23; Length 11;  
 Best Local Similarity 90.9%; Pred. No. 0.0075;  
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 EEVVPXGMSYS 11  
 |||||  
 DB 1 EEVVPXGMDYS 11  
 |||||  
 RESULT 13  
 ABB80528  
 ID ABB80528 standard; peptide; 11 AA.  
 XX  
 AC ABB80528;  
 XX  
 DT 08-OCT-2002 (first entry)  
 XX  
 DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #8.  
 XX  
 KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
 KW virucide.  
 XX  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Modified-site 1 /note= "N-terminal acetyl"  
 FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with  
 FT residue 7"  
 FT Misc-difference 8 /note= "D-form residue"  
 FT Modified-site 11 /note= "D-form residue"  
 FT Modified-site 11 /note= "C-terminal amide"  
 PN WO200208251-A2.  
 XX  
 PD 31-JAN-2002.  
 XX  
 PF 19-JUL-2001; 2001WO-US23169.  
 XX  
 PR 21-JUL-2000; 2000US-220101P.  
 XX  
 PA (CORV-) CORVAS INT INC.  
 XX  
 PI Lim-wilby M, Levy OE, Brunck TK;  
 XX  
 DR WPI; 2002-361643/39.  
 XX  
 PT Novel peptide compound having hepatitis C virus protease inhibitory  
 PT activity useful for treating disorders associated with hepatitis C  
 PT virus protease

PI Lim-wilby M, Levy OE, Brunck TK;  
 XX  
 DR WPI; 2002-361643/39.  
 XX  
 PT Novel peptide compound having hepatitis C virus protease inhibitory  
 PT activity useful for treating disorders associated with hepatitis C  
 PT virus protease  
 XX  
 PS Claim 17; Page 64; 69pp; English.  
 XX  
 CC The sequence represents a peptide compound of the invention having  
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
 CC invention are alpha-ketoamide peptide analogues. The peptides have  
 CC virucide activity, and are useful for treating and in the manufacture of  
 CC a medicament to treat disorders associated with HCV protease. A  
 CC pharmaceutical composition comprising the peptide as an active ingredient  
 CC is useful for treating disorders associated with hepatitis C virus.  
 XX  
 SQ Sequence 11 AA;  
 Query Match 88.5%; Score 46; DB 23; Length 11;  
 Best Local Similarity 90.9%; Pred. No. 0.0075;  
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 EEVVPXGMSYS 11  
 |||||  
 DB 1 EEVVPXGMDYS 11  
 |||||  
 RESULT 14  
 ABB80529  
 ID ABB80529 standard; peptide; 11 AA.  
 XX  
 AC ABB80529;  
 XX  
 DT 08-OCT-2002 (first entry)  
 XX  
 DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #9.  
 XX  
 KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
 KW virucide.  
 XX  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Modified-site 1 /note= "N-terminal acetyl"  
 FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with  
 FT residue 7"  
 FT Misc-difference 8 /note= "D-form residue"  
 FT Misc-difference 9 /note= "D-form residue"  
 FT Modified-site 11 /note= "D-form residue"  
 FT Modified-site 11 /note= "C-terminal amide"  
 PN WO200208251-A2.  
 XX  
 PD 31-JAN-2002.  
 XX  
 PF 19-JUL-2001; 2001WO-US23169.  
 XX  
 PR 21-JUL-2000; 2000US-220101P.  
 XX  
 PA (CORV-) CORVAS INT INC.  
 XX  
 PI Lim-wilby M, Levy OE, Brunck TK;  
 XX  
 DR WPI; 2002-361643/39.  
 XX  
 PT Novel peptide compound having hepatitis C virus protease inhibitory  
 PT activity useful for treating disorders associated with hepatitis C

PT virus protease -  
 PS Claim 17; Page 64; 69pp; English.  
 XX

CC The sequence represents a peptide compound of the invention having  
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
 CC invention are alpha-ketoamide peptide analogues. The peptides have  
 CC virucide activity, and are useful for treating and in the manufacture of  
 CC a medicament to treat disorders associated with HCV protease. A  
 CC pharmaceutical composition comprising the peptide as an active ingredient  
 CC is useful for treating disorders associated with hepatitis C virus.

XX SQ Sequence 11 AA;

Query Match 88.5%; Score 46; DB 23; Length 11;  
 Best Local Similarity 90.9%; Pred. No. 0.0075;  
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 EEVVPXGMSYS 11  
 |||||  
 Db 1 EEVVPXGMDYS 11

RESULT 15  
 ABB0561

ID ABB0561 standard; peptide; 11 AA.

XX AC ABB0561;

XX DT 08-OCT-2002 (first entry)

XX DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #41.

XX KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
 virucide.

XX OS Synthetic.

XX FH Key Location/Qualifiers

FT Modified-site 1 /note= "N-terminal acetyl"

FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with  
 residue 7"

FT Misc-difference 8 /note= "D-form residue"

FT Modified-site 8 /note= "Oxymethionine"

FT Modified-site 11 /note= "C-terminal amide"

XX WO200208251-A2.

XX PD 31-JAN-2002.

XX PF 19-JUL-2001; 2001WO-US23169.

XX PR 21-JUL-2000; 2000US-220101P.

XX PA (CORV-) CORVAS INT INC.

XX PI Lim-wilby M, Levy OE, Brunck TK;

XX XX WPI; 2002-361643/39.

XX Novel peptide compound having hepatitis C virus protease inhibitory  
 activity useful for treating disorders associated with hepatitis C  
 virus protease -  
 PS Claim 17; Page 65; 69pp; English.  
 XX

CC The sequence represents a peptide compound of the invention having  
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
 CC invention are alpha-ketoamide peptide analogues. The peptides have  
 CC virucide activity, and are useful for treating and in the manufacture of  
 CC a medicament to treat disorders associated with HCV protease. A  
 CC pharmaceutical composition comprising the peptide as an active ingredient  
 CC is useful for treating disorders associated with hepatitis C virus.

CC invention are alpha-ketoamide peptide analogues. The peptides have  
 CC virucide activity, and are useful for treating and in the manufacture of  
 CC a medicament to treat disorders associated with HCV protease. A  
 CC pharmaceutical composition comprising the peptide as an active ingredient  
 CC is useful for treating disorders associated with hepatitis C virus.

XX SQ Sequence 11 AA;

Query Match 88.5%; Score 46; DB 23; Length 11;  
 Best Local Similarity 90.9%; Pred. No. 0.0075;  
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 EEVVPXGMSYS 11  
 |||||  
 Db 1 EEVVPXGMDYS 11

RESULT 16  
 ABB0562

ID ABB0562 standard; peptide; 11 AA.

XX AC ABB0562;

XX DT 08-OCT-2002 (first entry)

XX DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #42.

XX KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
 virucide.

XX OS Synthetic.

XX FH Key Location/Qualifiers

FT Modified-site 1 /note= "N-terminal acetyl"

FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with  
 residue 7"

FT Misc-difference 8 /note= "D-form residue"

FT Modified-site 8 /note= "Oxymethionine"

FT Misc-difference 9 /note= "D-form residue"

FT Modified-site 11 /note= "C-terminal amide"

XX WO200208251-A2.

XX PD 31-JAN-2002.

XX PF 19-JUL-2001; 2001WO-US23169.

XX PR 21-JUL-2000; 2000US-220101P.

XX PA (CORV-) CORVAS INT INC.

XX PI Lim-wilby M, Levy OE, Brunck TK;

XX XX WPI; 2002-361643/39.

XX Novel peptide compound having hepatitis C virus protease inhibitory  
 activity useful for treating disorders associated with hepatitis C  
 virus protease -  
 PS Claim 17; Page 65; 69pp; English.  
 XX

CC The sequence represents a peptide compound of the invention having  
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
 CC invention are alpha-ketoamide peptide analogues. The peptides have  
 CC virucide activity, and are useful for treating and in the manufacture of  
 CC a medicament to treat disorders associated with HCV protease. A  
 CC pharmaceutical composition comprising the peptide as an active ingredient  
 CC is useful for treating disorders associated with hepatitis C virus.

CC is useful for treating disorders associated with hepatitis C virus.

XX Sequence 11 AA;

Query Match 88.5%; Score 46; DB 23; Length 11;

Best Local Similarity 90.9%; Pred. No. 0.0075; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXGMSYS 11

DB 1 EEVVPXGMDYS 11

#### RESULT 17

ID ABB80523 standard; peptide; 11 AA.

XX AC ABB80523;

XX DT 08-OCT-2002 (first entry)

XX DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #3.

XX KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;

XX OS virucide.

XX XX Synthetic.

XX FH Key Location/Qualifiers

FT Modified-site 1 /note= "N-terminal acetyl"

FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with

FT residue 7"

FT Misc-difference 9 /note= "D-form residue"

FT Modified-site 11 /note= "C-terminal amide"

FT WO200208251-A2.

PN 31-JAN-2002.

XX 19-JUL-2001; 2001WO-US23169.

XX 21-JUL-2000; 2000US-220101P.

XX (CORV-) CORVAS INT INC.

XX Lim-wilby M, Levy OE, Brunck TK;

XX WPI; 2002-361643/39.

XX Novel peptide compound having hepatitis C virus protease inhibitory

XX activity useful for treating disorders associated with hepatitis C

XX virus protease

XX Claim 17; Page 64; 69pp; English.

XX The sequence represents a peptide compound of the invention having

XX hepatitis C virus (HCV) protease inhibitory activity. The peptides of the

XX invention are alpha-ketoamide peptide analogues. The peptides have

XX virucide activity, and are useful for treating and in the manufacture of

XX a medicament to treat disorders associated with HCV protease. A

XX pharmaceutical composition comprising the peptide as an active ingredient

XX is useful for treating disorders associated with hepatitis C virus.

XX Sequence 11 AA;

Query Match 86.5%; Score 45; DB 23; Length 11;

Best Local Similarity 90.9%; Pred. No. 0.012;

Mismatches 1; Indels 0; Gaps 0;

XX Mismatches 10; Conservative 0;

QY 1 EEVVPXGMSYS 11

DB 1 EEVVPXGMDYS 11

#### RESULT 18

ID ABB80527 standard; peptide; 11 AA.

XX AC ABB80527;

XX DT 08-OCT-2002 (first entry)

XX DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #7.

XX KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;

XX OS virucide.

XX XX Synthetic.

XX FH Key Location/Qualifiers

FT Modified-site 1 /note= "N-terminal acetyl"

FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with

FT residue 7"

FT Misc-difference 8 /note= "D-form residue"

FT Modified-site 11 /note= "C-terminal amide"

FT WO200208251-A2.

PN 31-JAN-2002.

XX 19-JUL-2001; 2001WO-US23169.

XX 21-JUL-2000; 2000US-220101P.

XX (CORV-) CORVAS INT INC.

XX Lim-wilby M, Levy OE, Brunck TK;

XX WPI; 2002-361643/39.

XX Novel peptide compound having hepatitis C virus protease inhibitory

XX activity useful for treating disorders associated with hepatitis C

XX virus protease

XX Claim 17; Page 64; 69pp; English.

XX The sequence represents a peptide compound of the invention having

XX hepatitis C virus (HCV) protease inhibitory activity. The peptides of the

XX invention are alpha-ketoamide peptide analogues. The peptides have

XX virucide activity, and are useful for treating and in the manufacture of

XX a medicament to treat disorders associated with HCV protease. A

XX pharmaceutical composition comprising the peptide as an active ingredient

XX is useful for treating disorders associated with hepatitis C virus.

XX Sequence 11 AA;

Query Match 86.5%; Score 45; DB 23; Length 11;

Best Local Similarity 90.9%; Pred. No. 0.012;

Mismatches 1; Indels 0; Gaps 0;

XX Mismatches 10; Conservative 0;

QY 1 EEVVPXGMSYS 11

DB 1 EEVVPXGMDYS 11

#### RESULT 19

ID ABB80535 standard; peptide; 11 AA.

XX AC ABB80535;

XX DT 08-OCT-2002 (first entry)

XX DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #7.

XX KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;

XX OS virucide.

XX XX Synthetic.

XX FH Key Location/Qualifiers

FT Modified-site 1 /note= "N-terminal acetyl"

FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with

FT residue 7"

FT Misc-difference 8 /note= "D-form residue"

FT Modified-site 11 /note= "C-terminal amide"

FT WO200208251-A2.

PN 31-JAN-2002.

XX 19-JUL-2001; 2001WO-US23169.

XX 21-JUL-2000; 2000US-220101P.

XX (CORV-) CORVAS INT INC.

XX Lim-wilby M, Levy OE, Brunck TK;

XX WPI; 2002-361643/39.

XX Novel peptide compound having hepatitis C virus protease inhibitory

XX activity useful for treating disorders associated with hepatitis C

XX virus protease

XX Claim 17; Page 64; 69pp; English.

XX The sequence represents a peptide compound of the invention having

XX hepatitis C virus (HCV) protease inhibitory activity. The peptides of the

XX invention are alpha-ketoamide peptide analogues. The peptides have

XX virucide activity, and are useful for treating and in the manufacture of

XX a medicament to treat disorders associated with HCV protease. A

XX pharmaceutical composition comprising the peptide as an active ingredient

XX is useful for treating disorders associated with hepatitis C virus.

XX Sequence 11 AA;

Query Match 86.5%; Score 45; DB 23; Length 11;

Best Local Similarity 90.9%; Pred. No. 0.012;

Mismatches 1; Indels 0; Gaps 0;

XX Mismatches 10; Conservative 0;

QY 1 EEVVPXGMSYS 11

DB 1 EEVVPXGMDYS 11

OS	Synthetic.			
XX	Key	Location/Qualifiers		
XX	Modified-site	1	/note= "N-terminal acetyl"	
FT	Modified-site	6	/note= "Norvalyl carbonyl forming keto-amide linkage with residue 7"	
FT	Misc-difference	9	/note= "D-form residue"	
FT	Modified-site	11	/note= "C-terminal amide"	
XX	WO200208251-A2.			
XX	31-JAN-2002.			
XX	19-JUL-2001; 2001WO-US23169.			
XX	21-JUL-2000; 2000US-220101P.			
XX	(CORV-) CORVAS INT INC.			
XX	Lim-wilby M, Levy OE, Brunck TK;			
XX	WPI; 2002-361643/39.			
XX	Novel peptide compound having hepatitis C virus protease inhibitory activity useful for treating disorders associated with hepatitis C virus protease			
XX	Claim 17; Page 64; 69pp; English.			
XX	The sequence represents a peptide compound of the invention having hepatitis C virus (HCV) protease inhibitory activity. The peptides of the invention are alpha-ketoamide peptide analogues. The peptides have virucide activity, and are useful for treating and in the manufacture of a medicament to treat disorders associated with HCV protease. A pharmaceutical composition comprising the peptide as an active ingredient is useful for treating disorders associated with hepatitis C virus.			
XX	Sequence	11 AA;		
XX	Query Match	86.5%;	Score 45;	DB 23; Length 11;
XX	Best Local Similarity	90.9%;	Pred. No. 0.012;	1; Indels 0; Gaps 0;
XX	Matches 10;	Conservative 0;	Mismatches 1;	
QY	1	EEVVPXGMSYS 11		
DB	1	EEVVPXGQSYS 11		
XX	RESULT 21			
XX	ABB80539			
XX	ID	ABB80539	standard; peptide; 11 AA.	
XX	AC	ABB80539;		
XX	08-OCT-2002 (first entry)			
XX	Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #19.			
XX	Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide; virucide.			
XX	Synthetic.			
XX	Key	Location/Qualifiers		
XX	Modified-site	1	/note= "N-terminal acetyl"	
XX	Modified-site	6	/note= "Norvalyl carbonyl forming keto-amide linkage with residue 7"	

RESULT 21	
ABB80539	
ID	ABB80539 standard; peptide; 11 AA.
XX	
XX	ABB80539;
XX	
XX	08-OCT-2002 (first entry)
XX	
DE	Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #19.
XX	
XX	Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
KW	virucide.
KW	
XX	
XX	Synthetic.
XX	
XX	
Key	Location/Qualifiers
FT	Modified-site 1
FT	/note= "N-terminal acetyl"
FT	Modified-site 6
FT	/note= "Norvalyl carbonyl forming keto-amide linkage with
FT	residue 7"
FT	



XX Lim-wilby M, Levy OE, Brunck TK;  
 PI WPI; 2002-361643/39.  
 DR Novel peptide compound having hepatitis C virus protease inhibitory  
 XX activity useful for treating disorders associated with hepatitis C  
 PT virus protease -  
 PT  
 XX Claim 17; Page 65; 69pp; English.  
 PS The sequence represents a peptide compound of the invention having  
 XX hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
 CC invention are alpha-ketoamide peptide analogues. The peptides have  
 CC virucide activity, and are useful for treating and in the manufacture of  
 CC a medicament to treat disorders associated with HCV protease. A  
 CC pharmaceutical composition comprising the peptide as an active ingredient  
 CC is useful for treating disorders associated with hepatitis C virus.  
 XX Sequence 11 AA;  
 SQ

Query Match 86.5%; Score 45; DB 23; Length 11;  
 Best Local Similarity 90.9%; Pred. No. 0.012;  
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXGMSYS 11  
 DB 1 EEVVPXGMHYS 11

RESULT 24  
 ABB80560  
 ID ABB80560 standard; peptide; 11 AA.  
 XX AC ABB80560;  
 XX DT 08-OCT-2002 (first entry)  
 XX DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #40.  
 XX KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
 XX virucide.  
 XX OS Synthetic.  
 XX FH Key Location/Qualifiers  
 FT Modified-site 1 /note= "N-terminal acetyl"  
 FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with  
 FT residue 7"  
 FT Misc-difference 8 /note= "D-form residue"  
 FT Modified-site 8 /note= "Oxymethionine"  
 FT Misc-difference 9 /note= "D-form residue"  
 FT Modified-site 11 /note= "C-terminal amide"  
 FT WO200208251-A2.  
 XX PN 31-JAN-2002.  
 XX PD 19-JUL-2001; 2001WO-US23169.  
 XX PF 21-JUL-2000; 2000US-220101P.  
 XX PR (CORV-) CORVAS INT INC.  
 XX PA Lim-wilby M, Levy OE, Brunck TK;  
 XX PI WPI; 2002-361643/39.  
 XX DR Novel peptide compound having hepatitis C virus protease inhibitory  
 XX activity useful for treating disorders associated with hepatitis C  
 XX PT virus protease -  
 XX PS Claim 17; Page 65; 69pp; English.  
 XX CC The sequence represents a peptide compound of the invention having  
 XX hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
 CC invention are alpha-ketoamide peptide analogues. The peptides have  
 CC virucide activity, and are useful for treating and in the manufacture of  
 CC a medicament to treat disorders associated with HCV protease. A  
 CC pharmaceutical composition comprising the peptide as an active ingredient  
 CC is useful for treating disorders associated with hepatitis C virus.  
 XX Sequence 11 AA;  
 SQ

XX Novel peptide compound having hepatitis C virus protease inhibitory  
 PT activity useful for treating disorders associated with hepatitis C  
 PT virus protease -  
 XX Claim 17; Page 65; 69pp; English.  
 XX The sequence represents a peptide compound of the invention having  
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
 CC invention are alpha-ketoamide peptide analogues. The peptides have  
 CC virucide activity, and are useful for treating and in the manufacture of  
 CC a medicament to treat disorders associated with HCV protease. A  
 CC pharmaceutical composition comprising the peptide as an active ingredient  
 CC is useful for treating disorders associated with hepatitis C virus.  
 XX Sequence 11 AA;  
 SQ

Query Match 86.5%; Score 45; DB 23; Length 11;  
 Best Local Similarity 90.9%; Pred. No. 0.012;  
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXGMSYS 11  
 DB 1 EEVVPXGMHYS 11

RESULT 25  
 ABB80544  
 ID ABB80544 standard; peptide; 11 AA.  
 XX AC ABB80544;  
 XX DT 08-OCT-2002 (first entry)  
 XX DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #24.  
 XX KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
 XX virucide.  
 XX OS Synthetic.  
 XX FH Key Location/Qualifiers  
 FT Modified-site 1 /note= "N-terminal acetyl"  
 FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with  
 FT residue 7"  
 FT Modified-site 11 /note= "C-terminal amide"  
 FT WO200208251-A2.  
 XX PN 31-JAN-2002.  
 XX PD 19-JUL-2001; 2001WO-US23169.  
 XX PF 21-JUL-2000; 2000US-220101P.  
 XX PR (CORV-) CORVAS INT INC.  
 XX PA Lim-wilby M, Levy OE, Brunck TK;  
 XX PI WPI; 2002-361643/39.  
 XX DR Novel peptide compound having hepatitis C virus protease inhibitory  
 XX activity useful for treating disorders associated with hepatitis C  
 XX PT virus protease -  
 XX PS Claim 17; Page 65; 69pp; English.  
 XX CC The sequence represents a peptide compound of the invention having  
 XX hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
 CC invention are alpha-ketoamide peptide analogues. The peptides have  
 CC virucide activity, and are useful for treating and in the manufacture of  
 CC a medicament to treat disorders associated with HCV protease. A  
 CC pharmaceutical composition comprising the peptide as an active ingredient  
 CC is useful for treating disorders associated with hepatitis C virus.  
 XX Sequence 11 AA;  
 SQ



CC virucide activity, and are useful for treating and in the manufacture of  
CC a medicament to treat disorders associated with HCV protease. A  
CC pharmaceutical composition comprising the peptide as an active ingredient  
CC is useful for treating disorders associated with hepatitis C virus.  
XX

SQ Sequence 11 AA;

Query Match 84.6%; Score 44; DB 23; Length 11;  
Best Local Similarity 90.9%; Pred. No. 0.019;  
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXGMSYS 11  
|||  
Db 1 EEVVPXGTSYS 11

Search completed: June 10, 2003, 13:39:05  
Job time : 32.3571 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2003, 13:31:45 ; Search time 9.64286 Seconds  
(without alignments)  
33.564 Million cell updates/sec

Title: US-09-909-164-5  
Perfect score: 52  
Sequence: 1 EEVVPXGMSYS 11

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 75 summaries

Database :

Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/1/iaa/5A.COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B.COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A.COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B.COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PCFUS.COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	34	65.4	947	4	US-09-228-986-73
2	33	63.5	45	2	US-08-637-759B-236
3	33	63.5	45	3	US-08-871-355A-236
4	33	63.5	45	4	US-09-201-945-236
5	33	63.5	65	6	5177197-51
6	33	63.5	410	6	5177197-1
7	33	63.5	1394	6	5177197-30
8	32	61.5	10	4	US-09-357-952-66
9	32	61.5	10	4	US-09-521-650-66
10	32	61.5	10	4	US-09-168-888-66
11	32	61.5	102	2	US-08-580-988A-23
12	32	61.5	152	2	US-08-460-694-4
13	32	61.5	152	3	US-08-460-744-4
14	32	61.5	152	3	US-07-667-711B-4
15	32	61.5	173	1	US-08-193-977-7
16	32	61.5	189	2	US-08-464-517-21
17	32	61.5	189	2	US-08-246-361A-21
18	32	61.5	189	3	US-08-463-772-21
19	32	61.5	189	5	PCT-US93-05000-21
20	32	61.5	236	2	US-08-464-517-22
21	32	61.5	236	2	US-08-246-361A-22
22	32	61.5	236	3	US-08-463-772-22
23	32	61.5	236	5	PCT-US93-05000-22
24	32	61.5	280	2	US-08-464-517-6
25	32	61.5	280	3	US-08-463-772-6
26	32	61.5	289	2	US-08-246-361A-4
27	32	61.5	289	5	PCT-US93-05000-4

28	32	61.5	291	5	PCT-US93-05000-6
29	32	61.5	292	2	US-08-464-517-23
30	32	61.5	292	2	US-08-246-361A-6
31	32	61.5	292	2	US-08-246-361A-23
32	32	61.5	292	3	US-08-463-772-23
33	32	61.5	292	5	PCT-US93-05000-23
34	32	61.5	295	1	US-07-947-120-8
35	32	61.5	295	1	US-08-472-893A-8
36	32	61.5	295	2	US-08-460-694-2
37	32	61.5	295	2	US-08-464-517-19
38	32	61.5	295	2	US-08-464-517-20
39	32	61.5	295	2	US-08-246-361A-19
40	32	61.5	295	2	US-08-246-361A-20
41	32	61.5	295	3	US-08-463-772-19
42	32	61.5	295	3	US-08-463-772-20
43	32	61.5	295	3	US-08-460-744-2
44	32	61.5	295	3	US-07-667-711B-2
45	32	61.5	295	3	US-08-947-492-8
46	32	61.5	295	5	PCT-US93-05000-2
47	32	61.5	295	5	PCT-US93-05000-19
48	32	61.5	295	5	PCT-US93-05000-20
49	32	61.5	309	2	US-08-464-517-4
50	32	61.5	309	3	US-08-463-772-4
51	32	61.5	529	4	US-09-240-639-4
52	32	61.5	618	2	US-08-770-761A-3
53	32	61.5	647	2	US-08-770-761A-8
54	32	61.5	660	2	US-08-770-761A-2
55	32	61.5	662	2	US-08-770-761A-5
56	32	61.5	705	2	US-08-770-761A-7
57	32	61.5	819	2	US-08-464-517-7
58	32	61.5	819	2	US-08-246-361A-7
59	32	61.5	819	3	US-08-463-772-7
60	32	61.5	819	5	PCT-US93-05000-7
61	31	59.6	59	4	US-08-963-851-14
62	31	59.6	622	2	US-08-459-146-2
63	31	59.6	622	2	US-08-459-065-2
64	30	57.7	13	4	US-09-288-391-22
65	30	57.7	13	4	US-09-288-391-23
66	30	57.7	117	4	US-08-444-818-44
67	30	57.7	121	4	US-09-152-060-68
68	30	57.7	121	4	US-09-152-060-85
69	30	57.7	122	2	US-08-879-995A-1
70	30	57.7	122	3	US-09-215-096-1
71	30	57.7	132	4	US-08-444-818-52
72	30	57.7	159	2	US-08-844-086-4
73	30	57.7	159	3	US-09-018-211-4
74	30	57.7	211	5	PCT-US94-04174-18
75	30	57.7	241	3	US-08-834-776A-2

#### ALIGNMENTS

RESULT 1  
US-09-228-986-73  
; Sequence 73, Application US/09228986  
; Patent No. 6359198  
; GENERAL INFORMATION:  
; APPLICANT: Strabala, Timothy  
; APPLICANT: Nieuwenhuizen, Niels  
; TITLE OF INVENTION: Compositions Isolated from Plant Cells  
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signalling  
; FILE REFERENCE: 11000/1020  
; CURRENT APPLICATION NUMBER: US/09/228,986  
; CURRENT FILING DATE: 1999-01-12  
; NUMBER OF SEQ ID NOS: 130  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 73  
; LENGTH: 947  
; TYPE: PRT  
; ORGANISM: Pinus radiata  
US-09-228-986-73

Sequence 6, Appli  
Sequence 23, Appli  
Sequence 6, Appli  
Sequence 23, Appli  
Sequence 23, Appli  
Sequence 23, Appli  
Sequence 8, Appli  
Sequence 8, Appli  
Sequence 19, Appli  
Sequence 20, Appli  
Sequence 19, Appli  
Sequence 20, Appli  
Sequence 2, Appli  
Sequence 2, Appli  
Sequence 8, Appli  
Sequence 2, Appli  
Sequence 19, Appli  
Sequence 20, Appli  
Sequence 2, Appli  
Sequence 4, Appli  
Sequence 4, Appli  
Sequence 3, Appli  
Sequence 8, Appli  
Sequence 2, Appli  
Sequence 5, Appli  
Sequence 7, Appli  
Sequence 7, Appli  
Sequence 7, Appli  
Sequence 14, Appli  
Sequence 2, Appli  
Sequence 2, Appli  
Sequence 22, Appli  
Sequence 23, Appli  
Sequence 44, Appli  
Sequence 68, Appli  
Sequence 85, Appli  
Sequence 1, Appli  
Sequence 1, Appli  
Sequence 52, Appli  
Sequence 4, Appli  
Sequence 4, Appli  
Sequence 18, Appli  
Sequence 2, Appli

Query Match 65.4%; Score 34; DB 4; Length 947;  
Best Local Similarity 66.7%; Pred. No. 1e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMSYS 11  
| | | | |  
Db 686 VMPGSIYS 694

## RESULT 2

US-08-637-759B-236

; Sequence 236, Application US/08637759B

; Patent No. 5876931

; GENERAL INFORMATION:

; APPLICANT: David William Holden

; TITLE OF INVENTION: Identification of Genes

; NUMBER OF SEQUENCES: 501

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Patrea L. Pabst

; STREET: 2800 One Atlantic Center

; CITY: Atlanta

; STATE: Georgia

; COUNTRY: USA

; ZIP: 30309-3450

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/637,759B

; FILING DATE: 03-MAY-1996

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: PCT/GB95/02875

; FILING DATE: 11-DEC-1995

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Pabst, Patrea L.

; REGISTRATION NUMBER: 31,284

; REFERENCE/DOCKET NUMBER: RPMS 101

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (404) 873-8794

; TELEFAX: (404) 873-8795

; INFORMATION FOR SEQ ID NO: 236:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 45 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; HYPOTHETICAL: NO

US-08-637-759B-236

Query Match 63.5%; Score 33; DB 2; Length 45;  
Best Local Similarity 60.0%; Pred. No. 5.1;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10  
| | | | |  
Db 1 EEISPLGWSY 10

## RESULT 3

US-08-871-355A-236

; Sequence 236, Application US/08871355A

; Patent No. 6015669

; GENERAL INFORMATION:

; APPLICANT: David William Holden

; TITLE OF INVENTION: Identification of Genes

; NUMBER OF SEQUENCES: 501

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Patrea L. Pabst  
; STREET: 2800 One Atlantic Center  
; CITY: Atlanta  
; STATE: Georgia  
; COUNTRY: USA  
; ZIP: 30309-3450  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/871,355A  
; FILING DATE: 09-JUN-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/GB95/02875  
; FILING DATE: 11-DEC-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Pabst, Patrea L.  
; REGISTRATION NUMBER: 31,284  
; REFERENCE/DOCKET NUMBER: RPMS 101 CON  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (404) 873-8794  
; TELEFAX: (404) 873-8795  
; INFORMATION FOR SEQ ID NO: 236:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 45 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
; US-08-871-355A-236

Query Match 63.5%; Score 33; DB 3; Length 45;  
Best Local Similarity 60.0%; Pred. No. 5.1;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10  
| | | | |  
Db 1 EEISPLGWSY 10

## RESULT 4

US-09-201-945-236

; Sequence 236, Application US/09201945

; Patent No. 6342215

; GENERAL INFORMATION:

; APPLICANT: David William Holden

; TITLE OF INVENTION: Identification of Genes

; NUMBER OF SEQUENCES: 501

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Patrea L. Pabst

; STREET: 2800 One Atlantic Center

; CITY: Atlanta

; STATE: Georgia

; COUNTRY: USA

; ZIP: 30309-3450

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/201,945

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/637,759

;  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Padst, Patrea L.  
; REGISTRATION NUMBER: 31,284  
; REFERENCE/DOCKET NUMBER: RPMS 101  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (404) 873-8794  
; TELEFAX: (404) 873-8795  
; INFORMATION FOR SEQ ID NO: 236:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 45 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
; US-09-201-945-236

Query Match 63.5%; Score 33; DB 4; Length 45;  
Best Local Similarity 60.0%; Pred. No. 5.1;  
Matches 6; Conservative 1; Mismatches 3; Indels 3; Gaps 0;

QY 1 EEVVPXGMSY 10  
II: I I I I  
DB 1 EEISPLGWSY 10

RESULT 5  
517197-51  
; Patent No. 5177197  
; APPLICANT: KANZAKI, TETSUTO; OLOFSSON, ANDERS; MOREN, ANITA;  
; WERNSTEDT, CHRISTER; HELLMAN, ULF; MIYAZONO, KOHEI; CLAES-SON-WELSH,  
; LENA; HELDIN, CARL-HENRIK  
; TITLE OF INVENTION: ISOLATED NUCLEOTIDE SEQUENCE EXPRESSING  
; HUMAN TRANSFORMING GROWTH FACTOR-BETAL-BINDING PROTEIN  
; NUMBER OF SEQUENCES: 53  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/487,343  
; FILING DATE: 27-FEB-1990  
; SEQ ID NO: 51:  
; LENGTH: 65  
517197-51

Query Match 63.5%; Score 33; DB 6; Length 65;  
Best Local Similarity 45.5%; Pred. No. 7.8;  
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSYS 11  
II: I I I I  
DB 52 KEICPGMGYT 62

RESULT 6  
517197-1  
; Patent No. 5177197  
; APPLICANT: KANZAKI, TETSUTO; OLOFSSON, ANDERS; MOREN, ANITA;  
; WERNSTEDT, CHRISTER; HELLMAN, ULF; MIYAZONO, KOHEI; CLAES-SON-WELSH,  
; LENA; HELDIN, CARL-HENRIK  
; TITLE OF INVENTION: ISOLATED NUCLEOTIDE SEQUENCE EXPRESSING  
; HUMAN TRANSFORMING GROWTH FACTOR-BETAL-BINDING PROTEIN  
; NUMBER OF SEQUENCES: 53  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/487,343  
; FILING DATE: 27-FEB-1990  
; SEQ ID NO: 1:  
; LENGTH: 410  
517197-1

Query Match 63.5%; Score 33; DB 6; Length 410;  
Best Local Similarity 45.5%; Pred. No. 63;  
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSYS 11  
II: I I I I  
DB 399 KEICPGMGYT 409

RESULT 7  
517197-30  
; Patent No. 5177197  
; APPLICANT: KANZAKI, TETSUTO; OLOFSSON, ANDERS; MOREN, ANITA;  
; WERNSTEDT, CHRISTER; HELLMAN, ULF; MIYAZONO, KOHEI; CLAES-SON-WELSH,  
; LENA; HELDIN, CARL-HENRIK  
; TITLE OF INVENTION: ISOLATED NUCLEOTIDE SEQUENCE EXPRESSING  
; HUMAN TRANSFORMING GROWTH FACTOR-BETAL-BINDING PROTEIN  
; NUMBER OF SEQUENCES: 53  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/487,343  
; FILING DATE: 27-FEB-1990  
; SEQ ID NO: 30:  
; LENGTH: 1394  
517197-30

Query Match 63.5%; Score 33; DB 6; Length 1394;  
Best Local Similarity 45.5%; Pred. No. 2.6e+02;  
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSYS 11  
II: I I I I  
DB 399 KEICPGMGYT 409

RESULT 8  
US-09-357-952-66  
; Sequence 66, Application US/09357952  
; Patent No. 6248904  
; GENERAL INFORMATION:  
; APPLICANT: Zhang, Han-Zhong  
; APPLICANT: Cai, Sui Xiong  
; APPLICANT: Drewe, John A.  
; APPLICANT: Yang, Wu  
; TITLE OF INVENTION: No. 6248904el Fluorescence Dyes and Their Applications for W  
; TITLE OF INVENTION: Fluorescence Screening Assays for Caspases, Peptidases, Prot  
; FILE REFERENCE: 1735.0030001  
; CURRENT APPLICATION NUMBER: US/09/357,952  
; CURRENT FILING DATE: 1999-07-21  
; EARLIER APPLICATION NUMBER: US 60/093,642  
; EARLIER FILING DATE: 21-JUL-1998  
; NUMBER OF SEQ ID NOS: 139  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 66  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
; OTHER INFORMATION: Peptide  
US-09-357-952-66

Query Match 61.5%; Score 32; DB 4; Length 10;  
Best Local Similarity 50.0%; Pred. No. 1.5;  
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10  
II: I I I I  
DB 1 DDIVPCMSY 10

RESULT 9  
US-09-521-650-66  
; Sequence 66, Application US/09521650  
; Patent No. 6335429  
; GENERAL INFORMATION:  
; APPLICANT: Weber, Eckard

APPLICANT: Cai, Sui Xiong  
APPLICANT: Keana, John F.W.  
APPLICANT: Drewe, John A.  
APPLICANT: Zhang, Han-Zhong  
TITLE OF INVENTION: No. 6335429el Fluorogenic or Fluorescent Reporter Molecules and  
TITLE OF INVENTION: Their Applications for Whole-Cell Fluorescence  
TITLE OF INVENTION: Screening Assays for Caspases and Other Enzymes and the  
TITLE OF INVENTION: Use Thereof  
FILE REFERENCE: 1735.0290002  
CURRENT APPLICATION NUMBER: US/09/521,650  
CURRENT FILING DATE: 2000-03-08  
EARLIER APPLICATION NUMBER: 09/168,888  
EARLIER FILING DATE: 1998-10-09  
EARLIER APPLICATION NUMBER: US 60/061,582  
EARLIER FILING DATE: 1997-10-10  
EARLIER APPLICATION NUMBER: US 09/033,661  
EARLIER FILING DATE: 1998-03-03  
NUMBER OF SEQ ID NOS: 142  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 66  
LENGTH: 10  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-09-521-650-66

Query Match 61.5%; Score 32; DB 4; Length 10;  
Best Local Similarity 50.0%; Pred. No. 1.5;  
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10  
Db 1 DDIVPCMSY 10

RESULT 10  
US-09-168-888-66  
Sequence 66, Application US/09168888  
Patent No. 6342611  
GENERAL INFORMATION:  
APPLICANT: Weber, Eckard  
APPLICANT: Cai, Sui Xiong  
APPLICANT: Keana, John F.W.  
APPLICANT: Drewe, John A.  
APPLICANT: Zhang, Han-Zhong  
TITLE OF INVENTION: No. 6342611el Fluorogenic or Fluorescent Reporter Molecules and  
TITLE OF INVENTION: Their Applications for Whole-Cell Fluorescence  
TITLE OF INVENTION: Screening Assays for Caspases and Other Enzymes and the  
TITLE OF INVENTION: Use Thereof  
FILE REFERENCE: 1735.0290002  
CURRENT APPLICATION NUMBER: US/09/168,888  
CURRENT FILING DATE: 1998-10-09  
EARLIER APPLICATION NUMBER: US 60/061,582  
EARLIER FILING DATE: 1997-10-10  
EARLIER APPLICATION NUMBER: US 09/033,661  
EARLIER FILING DATE: 1998-03-03  
NUMBER OF SEQ ID NOS: 142  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 66  
LENGTH: 10  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-09-168-888-66

Query Match 61.5%; Score 32; DB 4; Length 10;  
Best Local Similarity 50.0%; Pred. No. 1.5;  
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10  
Db 1 DDIVPCMSY 10

RESULT 11  
US-08-580-988A-23  
Sequence 23, Application US/08580988A  
Patent No. 5856161  
GENERAL INFORMATION:  
APPLICANT: Aggarwal et al.  
TITLE OF INVENTION: Tumor Necrosis Factor  
TITLE OF INVENTION: Receptor-I-Associated Protein Kinase And Methods  
TITLE OF INVENTION: For Its Use  
NUMBER OF SEQUENCES: 27  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dr. Benjamin A. Adler  
STREET: 8011 Candle Lane  
CITY: Houston  
STATE: Texas  
COUNTRY: USA  
ZIP: 77071  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 1.44 Mb floppy disk  
COMPUTER: Apple Macintosh  
OPERATING SYSTEM: Macintosh  
SOFTWARE: Microsoft Word for Macintosh  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/580,988A  
FILING DATE: January 3, 1996  
CLASSIFICATION: 424  
PRIOR APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Benjamin Aaron Adler, Ph.D., J.D.  
REGISTRATION NUMBER: 35,423  
REFERENCE/DOCKET NUMBER: D5721CIP2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 713-777-2321  
TELEFAX: 713-777-6908  
INFORMATION FOR SEQ ID NO: 23:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 102 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE:  
DESCRIPTION: protein  
HYPOTHETICAL: no  
ANTI-SENSE: no  
FRAGMENT TYPE: internal  
ORIGINAL SOURCE:  
US-08-580-988A-23

Query Match 61.5%; Score 32; DB 2; Length 102;  
Best Local Similarity 60.0%; Pred. No. 21;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10  
Db 24 EEVFLPMY 33

RESULT 12  
US-08-460-694-4  
Sequence 4, Application US/08460694  
Patent No. 5856655  
GENERAL INFORMATION:  
APPLICANT: Arnold, Andrew  
TITLE OF INVENTION: PRAD1 Cyclin and its cDNA  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:

ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
STREET: 1100 New York Avenue, N.W., Suite 600  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/460,694  
FILING DATE: 02-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: McConathy, Evelyn H.  
REGISTRATION NUMBER: 35,279  
REFERENCE/DOCKET NUMBER: 0609.4070002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-2600  
TELEFAX: 202-371-2540  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 152 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-460-694-4

Query Match 61.5%; Score 32; DB 2; Length 152;  
Best Local Similarity 60.0%; Pred. No. 33;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10  
||| | | |  
Db 20 EEVFPPLAMNY 29

RESULT 13  
US-08-460-744-4  
Sequence 4, Application US/08460744  
Patent No. 6107541  
GENERAL INFORMATION:  
APPLICANT: Arnold, Andrew  
TITLE OF INVENTION: Pradi Cyclin and its cdna  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
STREET: 1100 New York Avenue, N.W., Suite 600  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/460,744  
FILING DATE: 02-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: McConathy, Evelyn H.  
REGISTRATION NUMBER: 35,279  
REFERENCE/DOCKET NUMBER: 0609.4070005  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-2600  
TELEFAX: 202-371-2540  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:

LENGTH: 152 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-460-744-4  
Query Match 61.5%; Score 32; DB 3; Length 152;  
Best Local Similarity 60.0%; Pred. No. 33;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
QY 1 EEVVPXGMSY 10  
||| | | |  
Db 20 EEVFPPLAMNY 29  
RESULT 14  
US-07-667-711B-4  
Sequence 4, Application US/07667711B  
Patent No. 6110700  
GENERAL INFORMATION:  
APPLICANT: ARNOLD, ANDREW  
TITLE OF INVENTION: Pradi Cyclin and its cdna  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
STREET: 1100 NEW YORK AVE., NW, SUITE 600  
CITY: WASHINGTON  
STATE: DC  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/667,711B  
FILING DATE: 11-MAR-1991  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: MCPHAIL, DONALD R.  
REGISTRATION NUMBER: 35,811  
REFERENCE/DOCKET NUMBER: 0609.4070000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 371-2600  
TELEFAX: (202) 371-2540  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 152 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: not relevant  
MOLECULE TYPE: peptide  
US-07-667-711B-4

Query Match 61.5%; Score 32; DB 3; Length 152;  
Best Local Similarity 60.0%; Pred. No. 33;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10  
||| | | |  
Db 20 EEVFPPLAMNY 29

RESULT 15  
US-08-193-977-7  
Sequence 7, Application US/08193977  
Patent No. 5625031  
GENERAL INFORMATION:  
APPLICANT: WEBSTER, KEVIN R.  
APPLICANT: COLEMAN, KEVIN G.  
TITLE OF INVENTION: PEPTIDE INHIBITORS OF THE P33CDK2 AND

;; TITLE OF INVENTION: P34CDC2 CELL CYCLE REGULATORY KINASES AND HUMAN  
;; TITLE OF INVENTION: PAPILLOMAVIRUS E7 ONCOPROTEIN  
;; NUMBER OF SEQUENCES: 34  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: REED & ROBINS  
;; STREET: 635 BRYANT STREET  
;; CITY: PALO ALTO  
;; STATE: CALIFORNIA  
;; COUNTRY: UNITED STATES OF AMERICA  
;; ZIP: 94301  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/193,977  
;; FILING DATE: 08-FEB-1994  
;; CLASSIFICATION: 530  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: ROBINS, ROBERTA L.  
;; REGISTRATION NUMBER: 33,208  
;; REFERENCE/DOCKET NUMBER: 5998-0016  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (415) 617-8999  
;; TELEFAX: (415) 327-3231  
;; INFORMATION FOR SEQ ID NO: 7:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 173 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
US-08-193-977-7

Query Match 61.5%; Score 32; DB 1; Length 173;  
Best Local Similarity 60.0%; Pred No. 38;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EVVVPXGMSY 10  
||| | | |  
DB 55 EEVFPPLAMNY 64

RESULT 16  
US-08-464-517-21  
;; Sequence 21, Application US/08464517  
;; Patent No. 5869640  
;; GENERAL INFORMATION:  
;; APPLICANT: BEACH, David H.  
;; TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO  
;; NUMBER OF SEQUENCES: 50  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: LAHIVE & COCKFIELD  
;; STREET: 60 State Street  
;; CITY: Boston  
;; STATE: MA  
;; COUNTRY: USA  
;; ZIP: 02109  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: ASCII(text)  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/464,517  
;; FILING DATE:  
;; CLASSIFICATION: 435  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/963,308  
;; FILING DATE: 16-OCT-1992  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 189 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single

;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/701,514  
;; FILING DATE: 16-MAY-1991  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Matthew P. Vincent  
;; REGISTRATION NUMBER: 36,709  
;; REFERENCE/DOCKET NUMBER: MII-004C  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (617) 227-7400  
;; TELEFAX: (617) 227-5941  
;; INFORMATION FOR SEQ ID NO: 21:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 189 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
US-08-464-517-21

Query Match 61.5%; Score 32; DB 2; Length 189;  
Best Local Similarity 60.0%; Pred No. 42;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EVVVPXGMSY 10  
||| | | |  
DB 74 EEVFPPLAMNY 83

RESULT 17  
US-08-246-361A-21  
;; Sequence 21, Application US/08246361A  
;; Patent No. 5998582  
;; GENERAL INFORMATION:  
;; APPLICANT: BEACH, David H.  
;; TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO  
;; NUMBER OF SEQUENCES: 50  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: LAHIVE & COCKFIELD  
;; STREET: 60 State Street  
;; CITY: Boston  
;; STATE: MA  
;; COUNTRY: USA  
;; ZIP: 02109  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: ASCII(text)  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/246,361A  
;; FILING DATE: 19-MAY-1994  
;; CLASSIFICATION: 435  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/963,308  
;; FILING DATE: 16-OCT-1992  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/888,178  
;; FILING DATE: 26-MAY-1992  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/701,514  
;; FILING DATE: 16-MAY-1991  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Matthew P. Vincent  
;; REGISTRATION NUMBER: 36,709  
;; REFERENCE/DOCKET NUMBER: MII-004C  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (617) 227-7400  
;; TELEFAX: (617) 227-5941  
;; INFORMATION FOR SEQ ID NO: 21:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 189 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single



TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-246-361A-21

Query Match 61.5%; Score 32; DB 2; Length 189;  
Best Local Similarity 60.0%; Pred. No. 42;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10  
Db 74 EEVFPPLAMNY 83

## RESULT 18

US-08-463-772-21  
Sequence 21, Application US/08463772  
Patent No. 6066501

GENERAL INFORMATION:  
APPLICANT: BEACH, David H.  
TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO  
NUMBER OF SEQUENCES: 50  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD  
STREET: 60 State Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII(text)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/463,772  
FILING DATE:

CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/963,308  
FILING DATE: 16-OCT-1992  
APPLICATION NUMBER: US 07/888,178  
FILING DATE: 26-MAY-1992

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/701,514  
FILING DATE: 16-MAY-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Matthew P. Vincent  
REGISTRATION NUMBER: 36,709  
REFERENCE/DOCKET NUMBER: MII-004C  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 227-7400  
TELEFAX: (617) 227-5941

INFORMATION FOR SEQ ID NO: 21:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 189 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-463-772-21

Query Match 61.5%; Score 32; DB 3; Length 189;  
Best Local Similarity 60.0%; Pred. No. 42;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10  
Db 74 EEVFPPLAMNY 83

## RESULT 19

PCT-US93-05000-21  
Sequence 21, Application PC/TUS9305000

GENERAL INFORMATION:  
APPLICANT: MITOTIX  
TITLE OF INVENTION: D-Type Cyclin and Uses Related Thereto  
NUMBER OF SEQUENCES: 42  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
STREET: Two Militia Drive  
CITY: Lexington  
STATE: Massachusetts  
COUNTRY: US  
ZIP: 02173

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US93/05000  
FILING DATE: 19930525  
CLASSIFICATION:

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/888,178  
FILING DATE: 26-MAY-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Granahan, Patricia  
REGISTRATION NUMBER: 32,227  
REFERENCE/DOCKET NUMBER: CSHL91-02A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-861-6240  
TELEFAX: 616-861-9540

INFORMATION FOR SEQ ID NO: 21:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 189 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
PCT-US93-05000-21

Query Match 61.5%; Score 32; DB 5; Length 189;  
Best Local Similarity 60.0%; Pred. No. 42;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10  
Db 74 EEVFPPLAMNY 83

## RESULT 20

US-08-464-517-22  
Sequence 22, Application US/08464517  
Patent No. 5869640

GENERAL INFORMATION:  
APPLICANT: BEACH, David H.  
TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO  
NUMBER OF SEQUENCES: 50  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD  
STREET: 60 State Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII(text)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/464,517  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/963,308

; FILING DATE: 16-OCT-1992  
; APPLICATION NUMBER: US 07/888,178  
; FILING DATE: 26-MAY-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/701,514  
; FILING DATE: 16-MAY-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Matthew P. Vincent  
; REGISTRATION NUMBER: 36,709  
; REFERENCE/DOCKET NUMBER: MII-004C  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 227-7400  
; TELEFAX: (617) 227-5941  
; INFORMATION FOR SEQ ID NO: 22:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 236 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-464-517-22

Query Match 61.5%; Score 32; DB 2; Length 236;  
Best Local Similarity 60.0%; Pred. No. 54;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10  
||| | | |  
Db 20 EEVFPPLANNY 29

## RESULT 21

US-08-246-361A-22  
; Sequence 22, Application US/08246361A  
; Patent No. 5998582  
; GENERAL INFORMATION:  
; APPLICANT: BEACH, David H.  
; TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO  
; NUMBER OF SEQUENCES: 50  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD  
; STREET: 60 State Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII(text)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/246,361A  
; FILING DATE: 19-MAY-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/963,308  
; FILING DATE: 16-OCT-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/888,178  
; FILING DATE: 26-MAY-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/701,514  
; FILING DATE: 16-MAY-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Matthew P. Vincent  
; REGISTRATION NUMBER: 36,709  
; REFERENCE/DOCKET NUMBER: MII-004C  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 227-7400  
; TELEFAX: (617) 227-5941  
; INFORMATION FOR SEQ ID NO: 22:  
; SEQUENCE CHARACTERISTICS:

; LENGTH: 236 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-246-361A-22

Query Match 61.5%; Score 32; DB 2; Length 236;  
Best Local Similarity 60.0%; Pred. No. 54;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10  
||| | | |  
Db 20 EEVFPPLANNY 29

## RESULT 22

US-08-463-772-22  
; Sequence 22, Application US/08463772  
; Patent No. 6066501  
; GENERAL INFORMATION:  
; APPLICANT: BEACH, David H.  
; TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO  
; NUMBER OF SEQUENCES: 50  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD  
; STREET: 60 State Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII(text)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/463,772  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/963,308  
; FILING DATE: 16-OCT-1992  
; APPLICATION NUMBER: US 07/888,178  
; FILING DATE: 26-MAY-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/701,514  
; FILING DATE: 16-MAY-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Matthew P. Vincent  
; REGISTRATION NUMBER: 36,709  
; REFERENCE/DOCKET NUMBER: MII-004C  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 227-7400  
; TELEFAX: (617) 227-5941  
; INFORMATION FOR SEQ ID NO: 22:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 236 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-463-772-22

Query Match 61.5%; Score 32; DB 3; Length 236;  
Best Local Similarity 60.0%; Pred. No. 54;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10  
||| | | |  
Db 20 EEVFPPLANNY 29

RESULT 23  
PCT-US93-05000-22  
; Sequence 22, Application PC/TUS9305000  
; GENERAL INFORMATION:  
; APPLICANT: MITORIX  
; TITLE OF INVENTION: D-Type Cyclin and Uses Related Thereto  
; NUMBER OF SEQUENCES: 42  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
; STREET: Two Militia Drive  
; CITY: Lexington  
; STATE: Massachusetts  
; COUNTRY: US  
; ZIP: 02173  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US93/05000  
; FILING DATE: 19930525  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/07/888,178  
; FILING DATE: 26-MAY-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Granahan, Patricia  
; REGISTRATION NUMBER: 32,227  
; REFERENCE/DOCKET NUMBER: CSHL91-02A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-861-6240  
; TELEFAX: 616-861-9540  
; INFORMATION FOR SEQ ID NO: 22:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 236 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
PCT-US93-05000-22

Query Match 61.5%; Score 32; DB 5; Length 236;  
Best Local Similarity 60.0%; Pred. No. 54;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10  
Db 20 EEVFFPLAMNY 29

RESULT 24  
US-08-464-517-6  
; Sequence 6, Application US/08464517  
; Patent No. 5869640  
; GENERAL INFORMATION:  
; APPLICANT: BEACH, David H.  
; TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO  
; NUMBER OF SEQUENCES: 50  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD  
; STREET: 60 State Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII(text)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/464,517  
; FILING DATE:

CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/963,308  
; FILING DATE: 16-OCT-1992  
; APPLICATION NUMBER: US 07/888,178  
; FILING DATE: 26-MAY-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/701,514  
; FILING DATE: 16-MAY-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Matthew P. Vincent  
; REGISTRATION NUMBER: 36,709  
; REFERENCE/DOCKET NUMBER: MII-004C  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 227-7400  
; TELEFAX: (617) 227-5941  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 280 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-464-517-6

Query Match 61.5%; Score 32; DB 2; Length 280;  
Best Local Similarity 60.0%; Pred. No. 65;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10  
Db 75 EEVFFPLAMNY 84

RESULT 25  
US-08-463-772-6  
; Sequence 6, Application US/08463772  
; Patent No. 6066501  
; GENERAL INFORMATION:  
; APPLICANT: BEACH, David H.  
; TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO  
; NUMBER OF SEQUENCES: 50  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD  
; STREET: 60 State Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII(text)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/463,772  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/963,308  
; FILING DATE: 16-OCT-1992  
; APPLICATION NUMBER: US 07/888,178  
; FILING DATE: 26-MAY-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/701,514  
; FILING DATE: 16-MAY-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Matthew P. Vincent  
; REGISTRATION NUMBER: 36,709  
; REFERENCE/DOCKET NUMBER: MII-004C  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 227-7400  
; TELEFAX: (617) 227-5941  
; INFORMATION FOR SEQ ID NO: 6:

; SEQUENCE CHARACTERISTICS:  
; LENGTH: 280 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-463-772-6

Query Match 61.5%; Score 32; DB 3; Length 280;  
Best Local Similarity 60.0%; Pred. NO. 65;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 EEVVPXGMSY 10  
||| | |  
Db 75 EEVFPAMNY 84

Search completed: June 10, 2003, 13:51:31  
Job time : 11.6429 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2003, 13:46:50 ; Search time 15 Seconds  
(without alignments)  
75.710 Million cell updates/sec

Title: US-09-909-164-5  
Perfect score: 52  
Sequence: 1 EHVFXGMSYS 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 392085 seqs, 103240269 residues

Total number of hits satisfying chosen parameters: 392085

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 75 summaries

- Database : Published Applications AA:\*
- 1: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
  - 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
  - 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
  - 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
  - 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
  - 6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
  - 7: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
  - 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
  - 9: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
  - 10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*
  - 11: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
  - 12: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep.\*
  - 13: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*
  - 14: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	38	73.1	3472	9 US-10-027-806-4	Sequence 4, Appli
2	38	73.1	3472	9 US-10-034-623-4	Sequence 4, Appli
3	38	73.1	3472	9 US-10-027-801-4	Sequence 4, Appli
4	34	65.4	947	9 US-10-101-4648-73	Sequence 73, Appl
5	33	63.5	426	9 US-10-214-766-43	Sequence 43, Appl
6	33	63.5	478	9 US-09-924-340-108	Sequence 108, App
7	33	63.5	478	9 US-09-992-600A-108	Sequence 108, App
8	33	63.5	478	9 US-09-746-783-184	Sequence 184, App
9	33	63.5	478	9 US-10-000-489-108	Sequence 108, App
10	33	63.5	478	9 US-10-000-986-108	Sequence 108, App
11	33	63.5	653	9 US-09-820-843A-26	Sequence 26, Appl
12	32	61.5	10	US-09-947-367-66	Sequence 66, Appl
13	32	61.5	254	10 US-09-778-927A-53	Sequence 53, Appl
14	32	61.5	289	9 US-10-024-066-2	Sequence 2, Appli
15	32	61.5	289	9 US-10-024-066-4	Sequence 4, Appli
16	32	61.5	289	10 US-09-919-497-54	Sequence 54, Appl
17	32	61.5	295	10 US-09-925-300-1061	Sequence 1061, Ap
18	32	61.5	529	10 US-09-923-304-4	Sequence 4, Appli
19	32	61.5	691	9 US-10-101-921-4	Sequence 4, Appli

20	61.5	691	10	US-09-925-731-2	Sequence 2, Appli
21	1377	1377	10	US-09-815-242-10384	Sequence 10384, A
22	2799	2799	9	US-10-151-736-4	Sequence 4, Appli
23	53	53	9	US-10-092-154-878	Sequence 878, App
24	59.6	59	10	US-09-764-847-878	Sequence 878, App
25	59.6	59	10	US-09-948-080-14	Sequence 14, Appl
26	59.6	161	9	US-09-738-626-5124	Sequence 5124, Ap
27	59.6	163	9	US-10-117-846-20	Sequence 20, Appl
28	59.6	192	9	US-09-986-480-171	Sequence 171, App
29	59.6	198	10	US-09-731-872-334	Sequence 334, App
30	59.6	223	9	US-09-738-626-6349	Sequence 6349, Ap
31	59.6	299	10	US-09-815-242-10697	Sequence 10697, A
32	59.6	381	9	US-09-975-139-5	Sequence 5, Appli
33	59.6	702	9	US-10-280-403-2	Sequence 2, Appli
34	59.6	702	10	US-09-907-479-2	Sequence 2, Appli
35	59.6	763	9	US-09-738-626-4454	Sequence 4454, Ap
36	59.6	1053	10	US-09-815-242-5136	Sequence 5136, Ap
37	59.6	1407	10	US-09-815-242-10439	Sequence 10439, A
38	59.6	1426	10	US-09-912-020-340	Sequence 340, App
39	57.7	7	9	US-09-909-062-1	Sequence 1, Appli
40	57.7	7	9	US-09-909-062-9	Sequence 9, Appli
41	57.7	7	9	US-09-909-062-130	Sequence 130, App
42	57.7	121	9	US-09-852-797-68	Sequence 68, Appl
43	57.7	121	9	US-09-852-797-85	Sequence 85, Appl
44	57.7	121	10	US-09-853-161-68	Sequence 68, Appl
45	57.7	121	10	US-09-853-161-85	Sequence 85, Appl
46	57.7	121	10	US-09-852-659A-68	Sequence 68, Appl
47	57.7	121	10	US-09-852-659A-85	Sequence 85, Appl
48	57.7	135	9	US-09-992-598-359	Sequence 359, App
49	57.7	135	9	US-09-989-293A-359	Sequence 359, App
50	57.7	135	9	US-09-989-735-359	Sequence 359, App
51	57.7	135	9	US-09-990-444-359	Sequence 359, App
52	57.7	135	9	US-09-990-730-359	Sequence 359, App
53	57.7	135	9	US-09-990-436-359	Sequence 359, App
54	57.7	135	9	US-09-991-181-359	Sequence 359, App
55	57.7	135	9	US-09-993-687-359	Sequence 359, App
56	57.7	135	9	US-09-989-734-359	Sequence 359, App
57	57.7	135	9	US-09-997-653-359	Sequence 359, App
58	57.7	135	9	US-10-174-590-444	Sequence 444, App
59	57.7	135	9	US-10-176-758-444	Sequence 444, App
60	57.7	135	9	US-10-175-737-444	Sequence 444, App
61	57.7	135	9	US-09-993-667-359	Sequence 359, App
62	57.7	135	9	US-10-173-706-444	Sequence 444, App
63	57.7	135	9	US-10-175-738-444	Sequence 444, App
64	57.7	135	9	US-10-175-752-444	Sequence 444, App
65	57.7	135	9	US-10-176-482-444	Sequence 444, App
66	57.7	135	9	US-10-176-757-444	Sequence 444, App
67	57.7	135	9	US-10-176-913-444	Sequence 444, App
68	57.7	135	9	US-10-180-552-444	Sequence 444, App
69	57.7	135	9	US-10-180-557-444	Sequence 444, App
70	57.7	135	9	US-09-990-438-359	Sequence 359, App
71	57.7	135	9	US-09-990-562-359	Sequence 359, App
72	57.7	135	9	US-09-997-428-359	Sequence 359, App
73	57.7	135	9	US-09-997-666-359	Sequence 359, App
74	57.7	135	9	US-10-173-700-444	Sequence 444, App
75	57.7	135	9	US-10-174-572-444	Sequence 444, App

ALIGNMENTS

RESULT 1  
US-10-027-806-4  
; Sequence 4, Application US/10027806  
; Patent No. US20020160476A1  
; GENERAL INFORMATION:  
; APPLICANT: Swanson, Ronald V.  
; APPLICANT: Feldman, Robert A.  
; APPLICANT: Schleper, Christa  
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM CENARCHAUM SYMBIOSUM  
; FILE REFERENCE: DORP 002A  
; CURRENT APPLICATION NUMBER: US/10/027,806  
; CURRENT FILING DATE: 2001-12-21

us-09-909-164-5.rapb

Wed Jun 11 15:44:03 2003

Query Match 73.1%; Score 38; DB 9; Length 3472;  
Best Local Similarity 54.5%; Pred. No. 1.2e+02;  
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;  
NUMBER OF SEQ ID NOS: 123  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 4  
LENGTH: 3472  
TYPE: PRT  
ORGANISM: Cenarchaeum symbiosum  
US-10-027-806-4

Query Match 73.1%; Score 38; DB 9; Length 3472;  
Best Local Similarity 54.5%; Pred. No. 1.2e+02;  
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXGMSYS 11  
|:|:| |:|:|  
Db 2294 EDVIPRGISFS 2304

RESULT 2  
US-10-034-623-4  
Sequence 4, Application US/10034623  
Publication No. US20020198365A1  
GENERAL INFORMATION:  
APPLICANT: Swanson, Ronald V.  
APPLICANT: Feldman, Robert A.  
APPLICANT: Schleper, Christa  
TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM CENARCHAEUM SYMBIOSUM  
FILE REFERENCE: DCRP.002A  
CURRENT APPLICATION NUMBER: US/10/034,623  
CURRENT FILING DATE: 2001-12-21  
PRIOR APPLICATION NUMBER: 09/408,020  
PRIOR FILING DATE: 1999-09-29  
PRIOR APPLICATION NUMBER: 60/102,294  
PRIOR FILING DATE: 1998-09-29  
NUMBER OF SEQ ID NOS: 123  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 4  
LENGTH: 3472  
TYPE: PRT  
ORGANISM: Cenarchaeum symbiosum  
US-10-034-623-4

Query Match 73.1%; Score 38; DB 9; Length 3472;  
Best Local Similarity 54.5%; Pred. No. 1.2e+02;  
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXGMSYS 11  
|:|:| |:|:|  
Db 2294 EDVIPRGISFS 2304

RESULT 3  
US-10-027-801-4  
Sequence 4, Application US/10027801  
Publication No. US20030054364A1  
GENERAL INFORMATION:  
APPLICANT: Swanson, Ronald V.  
APPLICANT: Feldman, Robert A.  
APPLICANT: Schleper, Christa  
TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM CENARCHAEUM SYMBIOSUM  
FILE REFERENCE: DCRP.002A  
CURRENT APPLICATION NUMBER: US/10/027,801  
CURRENT FILING DATE: 2001-12-21  
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/408,020  
PRIOR FILING DATE: EARLIER FILING DATE: 1999-09-29  
NUMBER OF SEQ ID NOS: 123  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 4  
LENGTH: 3472  
TYPE: PRT  
ORGANISM: Cenarchaeum symbiosum  
US-10-027-801-4

Query Match 73.1%; Score 38; DB 9; Length 3472;  
Best Local Similarity 54.5%; Pred. No. 1.2e+02;  
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXGMSYS 11  
|:|:| |:|:|  
Db 2294 EDVIPRGISFS 2304

RESULT 4  
US-10-101-464A-73  
Sequence 73, Application US/10101464A  
Publication No. US20030046728A1  
GENERAL INFORMATION:  
APPLICANT: Strabala, Timothy  
APPLICANT: Nieuwenhuizen, Nicolaas  
APPLICANT: Higgins, Colleen M.  
TITLE OF INVENTION: Compositions Isolated from Plant Cells  
TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signalling  
FILE REFERENCE: 11000.1020c2  
CURRENT APPLICATION NUMBER: US/10/101,464A  
CURRENT FILING DATE: 2002-03-18  
PRIOR APPLICATION NUMBER: 09/704,302  
PRIOR FILING DATE: 2000-11-01  
PRIOR APPLICATION NUMBER: 09/228,986  
PRIOR FILING DATE: 1999-01-12  
PRIOR APPLICATION NUMBER: 60/162,866  
PRIOR FILING DATE: 1999-11-01  
PRIOR APPLICATION NUMBER: PCT/US00/00724  
PRIOR FILING DATE: 2000-01-11  
NUMBER OF SEQ ID NOS: 989  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 73  
LENGTH: 947  
TYPE: PRT  
ORGANISM: Pinus radiata  
US-10-101-464A-73

Query Match 65.4%; Score 34; DB 9; Length 947;  
Best Local Similarity 66.7%; Pred. No. 1.8e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMSYS 11  
|:|:| |:|:|  
Db 686 VMPXGMSYS 694

RESULT 5  
US-10-214-766-43  
Sequence 43, Application US/10214766  
Publication No. US20030084473A1  
GENERAL INFORMATION:  
APPLICANT: Gocal, Greg  
TITLE OF INVENTION: NON-TRANSGENIC HERBICIDE RESISTANT PLANTS  
FILE REFERENCE: CA1138  
CURRENT APPLICATION NUMBER: US/10/214,766  
CURRENT FILING DATE: 2002-08-09  
PRIOR APPLICATION NUMBER: US 60/311,734  
PRIOR FILING DATE: 2001-08-09  
NUMBER OF SEQ ID NOS: 53  
SOFTWARE: Patent in version 3.2  
SEQ ID NO 43  
LENGTH: 426  
TYPE: PRT  
ORGANISM: Vibrio cholerae  
US-10-214-766-43

Query Match 63.5%; Score 33; DB 9; Length 426;  
Best Local Similarity 60.0%; Pred. No. 1.2e+02;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSYS 10

Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Db 223 EFVIPAGQSY 232  
|||  
2 EVVPXGMSYS 11  
|||  
239 EVAPAGASYN 248

Db 223 EFVIPAGQSY 232  
|||  
2 EVVPXGMSYS 11  
|||  
239 EVAPAGASYN 248

Db 223 EFVIPAGQSY 232  
|||  
2 EVVPXGMSYS 11  
|||  
239 EVAPAGASYN 248

Db 223 EFVIPAGQSY 232  
|||  
2 EVVPXGMSYS 11  
|||  
239 EVAPAGASYN 248

RESULT 8  
US-09-746-783-184  
; Sequence 184, Application US/09746783  
; Publication No. US20030044935A1  
; GENERAL INFORMATION:  
; APPLICANT: Jacobs, Kenneth  
; McCoy, John M.  
; LaVallie, Edward R.  
; Racie, Lisa A.  
; Treacy, Maurice  
; Spaulding, Vikki  
; Agostino, Michael J.  
; Howes, Steven H.  
; Pechtel, Kim

RESULT 8  
US-09-746-783-184  
; Sequence 184, Application US/09746783  
; Publication No. US20030044935A1  
; GENERAL INFORMATION:  
; APPLICANT: Jacobs, Kenneth  
; McCoy, John M.  
; LaVallie, Edward R.  
; Racie, Lisa A.  
; Treacy, Maurice  
; Spaulding, Vikki  
; Agostino, Michael J.  
; Howes, Steven H.  
; Pechtel, Kim

RESULT 8  
US-09-746-783-184  
; Sequence 184, Application US/09746783  
; Publication No. US20030044935A1  
; GENERAL INFORMATION:  
; APPLICANT: Jacobs, Kenneth  
; McCoy, John M.  
; LaVallie, Edward R.  
; Racie, Lisa A.  
; Treacy, Maurice  
; Spaulding, Vikki  
; Agostino, Michael J.  
; Howes, Steven H.  
; Pechtel, Kim

RESULT 8  
US-09-746-783-184  
; Sequence 184, Application US/09746783  
; Publication No. US20030044935A1  
; GENERAL INFORMATION:  
; APPLICANT: Jacobs, Kenneth  
; McCoy, John M.  
; LaVallie, Edward R.  
; Racie, Lisa A.  
; Treacy, Maurice  
; Spaulding, Vikki  
; Agostino, Michael J.  
; Howes, Steven H.  
; Pechtel, Kim

TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES  
NUMBER OF SEQUENCES: 231  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genetics Institute, Inc.  
STREET: 87 Cambridgepark Drive  
CITY: Cambridge  
STATE: MA  
COUNTRY: U.S.A.  
ZIP: 02140  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/746,783  
FILING DATE: 21-Dec-2000  
CLASSIFICATION: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Milasincic, Debra J.  
REGISTRATION NUMBER: 46,931  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 227-7400  
TELEFAX: (617) 742-4214  
INFORMATION FOR SEQ ID NO: 184:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 478 amino acids  
TYPE: amino acid  
STRANDEDNESS: <Unknown>  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 184:

TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES  
NUMBER OF SEQUENCES: 231  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genetics Institute, Inc.  
STREET: 87 Cambridgepark Drive  
CITY: Cambridge  
STATE: MA  
COUNTRY: U.S.A.  
ZIP: 02140  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/746,783  
FILING DATE: 21-Dec-2000  
CLASSIFICATION: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Milasincic, Debra J.  
REGISTRATION NUMBER: 46,931  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 227-7400  
TELEFAX: (617) 742-4214  
INFORMATION FOR SEQ ID NO: 184:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 478 amino acids  
TYPE: amino acid  
STRANDEDNESS: <Unknown>  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 184:

TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES  
NUMBER OF SEQUENCES: 231  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genetics Institute, Inc.  
STREET: 87 Cambridgepark Drive  
CITY: Cambridge  
STATE: MA  
COUNTRY: U.S.A.  
ZIP: 02140  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/746,783  
FILING DATE: 21-Dec-2000  
CLASSIFICATION: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Milasincic, Debra J.  
REGISTRATION NUMBER: 46,931  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 227-7400  
TELEFAX: (617) 742-4214  
INFORMATION FOR SEQ ID NO: 184:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 478 amino acids  
TYPE: amino acid  
STRANDEDNESS: <Unknown>  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 184:

TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES  
NUMBER OF SEQUENCES: 231  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genetics Institute, Inc.  
STREET: 87 Cambridgepark Drive  
CITY: Cambridge  
STATE: MA  
COUNTRY: U.S.A.  
ZIP: 02140  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/746,783  
FILING DATE: 21-Dec-2000  
CLASSIFICATION: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Milasincic, Debra J.  
REGISTRATION NUMBER: 46,931  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 227-7400  
TELEFAX: (617) 742-4214  
INFORMATION FOR SEQ ID NO: 184:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 478 amino acids  
TYPE: amino acid  
STRANDEDNESS: <Unknown>  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 184:

Query Match 63.5%; Score 33; DB 9; Length 478;  
Best Local Similarity 60.0%; Pred. No. 1.4e+02;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Query Match 63.5%; Score 33; DB 9; Length 478;  
Best Local Similarity 60.0%; Pred. No. 1.4e+02;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Query Match 63.5%; Score 33; DB 9; Length 478;  
Best Local Similarity 60.0%; Pred. No. 1.4e+02;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Query Match 63.5%; Score 33; DB 9; Length 478;  
Best Local Similarity 60.0%; Pred. No. 1.4e+02;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

RESULT 7  
US-09-992-600A-108  
; Sequence 108, Application US/09992600A  
; Publication No. US20030027161A1  
; GENERAL INFORMATION:  
; APPLICANT: Benjamin, Stephane  
; Tanaka, Hiroaki  
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF  
; FILE REFERENCE: 91.US4.DIV  
; CURRENT APPLICATION NUMBER: US/09/992,600A  
; CURRENT FILING DATE: 2001-11-13  
; PRIOR APPLICATION NUMBER: US 09/924,340  
; PRIOR FILING DATE: 2001-08-06  
; PRIOR APPLICATION NUMBER: PCT/IB01/01715  
; PRIOR FILING DATE: 2001-08-06  
; PRIOR APPLICATION NUMBER: US 60/305,456  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: US 60/302,277  
; PRIOR FILING DATE: 2001-06-29  
; PRIOR APPLICATION NUMBER: US 60/298,698  
; PRIOR FILING DATE: 2001-06-15  
; PRIOR APPLICATION NUMBER: US 60/293,574  
; PRIOR FILING DATE: 2001-05-25  
; NUMBER OF SEQ ID NOS: 114  
; SOFTWARE: JPatent  
; SEQ ID NO 108  
; LENGTH: 478  
; TYPE: PPT  
; ORGANISM: Homo sapiens  
US-09-992-600A-108

RESULT 7  
US-09-992-600A-108  
; Sequence 108, Application US/09992600A  
; Publication No. US20030027161A1  
; GENERAL INFORMATION:  
; APPLICANT: Benjamin, Stephane  
; Tanaka, Hiroaki  
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF  
; FILE REFERENCE: 91.US4.DIV  
; CURRENT APPLICATION NUMBER: US/09/992,600A  
; CURRENT FILING DATE: 2001-11-13  
; PRIOR APPLICATION NUMBER: US 09/924,340  
; PRIOR FILING DATE: 2001-08-06  
; PRIOR APPLICATION NUMBER: PCT/IB01/01715  
; PRIOR FILING DATE: 2001-08-06  
; PRIOR APPLICATION NUMBER: US 60/305,456  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: US 60/302,277  
; PRIOR FILING DATE: 2001-06-29  
; PRIOR APPLICATION NUMBER: US 60/298,698  
; PRIOR FILING DATE: 2001-06-15  
; PRIOR APPLICATION NUMBER: US 60/293,574  
; PRIOR FILING DATE: 2001-05-25  
; NUMBER OF SEQ ID NOS: 114  
; SOFTWARE: JPatent  
; SEQ ID NO 108  
; LENGTH: 478  
; TYPE: PPT  
; ORGANISM: Homo sapiens  
US-09-992-600A-108

RESULT 7  
US-09-992-600A-108  
; Sequence 108, Application US/09992600A  
; Publication No. US20030027161A1  
; GENERAL INFORMATION:  
; APPLICANT: Benjamin, Stephane  
; Tanaka, Hiroaki  
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF  
; FILE REFERENCE: 91.US4.DIV  
; CURRENT APPLICATION NUMBER: US/09/992,600A  
; CURRENT FILING DATE: 2001-11-13  
; PRIOR APPLICATION NUMBER: US 09/924,340  
; PRIOR FILING DATE: 2001-08-06  
; PRIOR APPLICATION NUMBER: PCT/IB01/01715  
; PRIOR FILING DATE: 2001-08-06  
; PRIOR APPLICATION NUMBER: US 60/305,456  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: US 60/302,277  
; PRIOR FILING DATE: 2001-06-29  
; PRIOR APPLICATION NUMBER: US 60/298,698  
; PRIOR FILING DATE: 2001-06-15  
; PRIOR APPLICATION NUMBER: US 60/293,574  
; PRIOR FILING DATE: 2001-05-25  
; NUMBER OF SEQ ID NOS: 114  
; SOFTWARE: JPatent  
; SEQ ID NO 108  
; LENGTH: 478  
; TYPE: PPT  
; ORGANISM: Homo sapiens  
US-09-992-600A-108

RESULT 7  
US-09-992-600A-108  
; Sequence 108, Application US/09992600A  
; Publication No. US20030027161A1  
; GENERAL INFORMATION:  
; APPLICANT: Benjamin, Stephane  
; Tanaka, Hiroaki  
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF  
; FILE REFERENCE: 91.US4.DIV  
; CURRENT APPLICATION NUMBER: US/09/992,600A  
; CURRENT FILING DATE: 2001-11-13  
; PRIOR APPLICATION NUMBER: US 09/924,340  
; PRIOR FILING DATE: 2001-08-06  
; PRIOR APPLICATION NUMBER: PCT/IB01/01715  
; PRIOR FILING DATE: 2001-08-06  
; PRIOR APPLICATION NUMBER: US 60/305,456  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: US 60/302,277  
; PRIOR FILING DATE: 2001-06-29  
; PRIOR APPLICATION NUMBER: US 60/298,698  
; PRIOR FILING DATE: 2001-06-15  
; PRIOR APPLICATION NUMBER: US 60/293,574  
; PRIOR FILING DATE: 2001-05-25  
; NUMBER OF SEQ ID NOS: 114  
; SOFTWARE: JPatent  
; SEQ ID NO 108  
; LENGTH: 478  
; TYPE: PPT  
; ORGANISM: Homo sapiens  
US-09-992-600A-108

Query Match 63.5%; Score 33; DB 9; Length 478;  
Best Local Similarity 60.0%; Pred. No. 1.4e+02;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Query Match 63.5%; Score 33; DB 9; Length 478;  
Best Local Similarity 60.0%; Pred. No. 1.4e+02;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Query Match 63.5%; Score 33; DB 9; Length 478;  
Best Local Similarity 60.0%; Pred. No. 1.4e+02;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Query Match 63.5%; Score 33; DB 9; Length 478;  
Best Local Similarity 60.0%; Pred. No. 1.4e+02;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

RESULT 9  
US-10-000-489-108  
; Sequence 108, Application US/10000489  
; Publication No. US20030092011A1  
; GENERAL INFORMATION:  
; APPLICANT: Benjamin, Stephane  
; Tanaka, Hiroaki  
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF

RESULT 9  
US-10-000-489-108  
; Sequence 108, Application US/10000489  
; Publication No. US20030092011A1  
; GENERAL INFORMATION:  
; APPLICANT: Benjamin, Stephane  
; Tanaka, Hiroaki  
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF

RESULT 9  
US-10-000-489-108  
; Sequence 108, Application US/10000489  
; Publication No. US20030092011A1  
; GENERAL INFORMATION:  
; APPLICANT: Benjamin, Stephane  
; Tanaka, Hiroaki  
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF

RESULT 9  
US-10-000-489-108  
; Sequence 108, Application US/10000489  
; Publication No. US20030092011A1  
; GENERAL INFORMATION:  
; APPLICANT: Benjamin, Stephane  
; Tanaka, Hiroaki  
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF

; FILE REFERENCE: 91.US6.DIV  
; CURRENT APPLICATION NUMBER: US/10/000,489  
; CURRENT FILING DATE: 2001-11-14  
; PRIOR APPLICATION NUMBER: US 09/924,340  
; PRIOR FILING DATE: 2001-08-06  
; PRIOR APPLICATION NUMBER: PCT/IB01/01715  
; PRIOR FILING DATE: 2001-08-06  
; PRIOR APPLICATION NUMBER: US 60/305,456  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: US 60/302,277  
; PRIOR FILING DATE: 2001-06-29  
; PRIOR APPLICATION NUMBER: US 60/298,698  
; PRIOR FILING DATE: 2001-06-15  
; PRIOR APPLICATION NUMBER: US 60/293,574  
; PRIOR FILING DATE: 2001-05-25  
; NUMBER OF SEQ ID NOS: 112  
; SOFTWARE: JPatent  
; SEQ ID NO 108  
; LENGTH: 478  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-000-489-108

Query Match 63.5%; Score 33; DB 9; Length 478;  
Best Local Similarity 60.0%; Pred. No. 1.4e+02;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 EVVPXGMSYS 11  
Db 239 EVAPAGASYN 248  
|||||

RESULT 10  
US-10-000-986-108  
; Sequence 108, Application US/10000986  
; Publication No. US20030096247A1  
; GENERAL INFORMATION:  
; APPLICANT: Benjanin, Stephane  
; APPLICANT: Tanaka, Hiroaki  
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF  
; FILE REFERENCE: 91.US9.DIV  
; CURRENT APPLICATION NUMBER: US/10/000,986  
; CURRENT FILING DATE: 2001-11-14  
; PRIOR APPLICATION NUMBER: US 09/924,340  
; PRIOR FILING DATE: 2001-08-06  
; PRIOR APPLICATION NUMBER: PCT/IB01/01715  
; PRIOR FILING DATE: 2001-08-06  
; PRIOR APPLICATION NUMBER: US 60/305,456  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: US 60/302,277  
; PRIOR FILING DATE: 2001-06-29  
; PRIOR APPLICATION NUMBER: US 60/298,698  
; PRIOR FILING DATE: 2001-06-15  
; PRIOR APPLICATION NUMBER: US 60/293,574  
; PRIOR FILING DATE: 2001-05-25  
; NUMBER OF SEQ ID NOS: 112  
; SOFTWARE: JPatent  
; SEQ ID NO 108  
; LENGTH: 478  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-000-986-108

Query Match 63.5%; Score 33; DB 9; Length 478;  
Best Local Similarity 60.0%; Pred. No. 1.4e+02;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 EVVPXGMSYS 11  
Db 239 EVAPAGASYN 248  
|||||

RESULT 11

US-09-820-843A-26  
; Sequence 26, Application US/09820843A  
; Publication No. US20030039963A1  
; GENERAL INFORMATION:  
; APPLICANT: Council of Scientific and Industrial Research  
; TITLE OF INVENTION: A COMPUTATIONAL METHOD FOR THE IDENTIFICATION OF CANDIDATE P  
; TITLE OF INVENTION: USEFUL AS ANTI-INFECTIVES  
; FILE REFERENCE: Q63915  
; CURRENT APPLICATION NUMBER: US/09/820,843A  
; CURRENT FILING DATE: 2001-03-30  
; NUMBER OF SEQ ID NOS: 118  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 26  
; LENGTH: 653  
; TYPE: PRT  
; ORGANISM: Vibrio cholerae  
; FEATURE:  
; NAME/KEY: misc.feature  
; OTHER INFORMATION: Iron(III) ABC transporter, permease protein  
; NAME/KEY: misc.feature  
; OTHER INFORMATION: g1|9654609  
US-09-820-843A-26

Query Match 63.5%; Score 33; DB 9; Length 653;  
Best Local Similarity 66.7%; Pred. No. 1.9e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXGMS 9  
Db 300 EEVVPXGMS 308  
|||||

RESULT 12  
US-09-947-387-66  
; Sequence 66, Application US/09947387  
; Patent No. US20020150885A1  
; GENERAL INFORMATION:  
; APPLICANT: Weber, Eckard  
; APPLICANT: Cai, Sui Xiong  
; APPLICANT: Keana, John F.W.  
; APPLICANT: Drewe, John A.  
; APPLICANT: Zhang, Han-Zhong  
; TITLE OF INVENTION: No. US20020150885A1el Fluorogenic or Fluorescent Reporter Mol  
; TITLE OF INVENTION: Their Applications for Whole-Cell Fluorescence  
; TITLE OF INVENTION: Screening Assays for Caspases and Other Enzymes and the  
; FILE REFERENCE: 1735.0290005  
; CURRENT APPLICATION NUMBER: US/09/947,387  
; CURRENT FILING DATE: 2001-09-07  
; PRIOR APPLICATION NUMBER: US 60/061,582  
; PRIOR FILING DATE: 1997-10-10  
; PRIOR APPLICATION NUMBER: US 60/145,746  
; PRIOR FILING DATE: 1998-03-03  
; PRIOR APPLICATION NUMBER: US 09/168,888  
; PRIOR FILING DATE: 1998-10-09  
; NUMBER OF SEQ ID NOS: 142  
; SOFTWARE: PatentIn ver. 2.0  
; SEQ ID NO 66  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
; OTHER INFORMATION: Peptide  
US-09-947-387-66

Query Match 61.5%; Score 32; DB 10; Length 10;  
Best Local Similarity 50.0%; Pred. No. 3.5;  
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10  
Db 1 DDIVPCMSY 10  
:::|



## RESULT 13

US-09-778-927A-53

; Sequence 53, Application US/09778927A

; Patent No. US20020068342A1

; GENERAL INFORMATION:

; APPLICANT: KHOSRAVI, Rami et al.

; TITLE OF INVENTION: NOVEL NUCLEIC ACID AND AMINO ACID SEQUENCES AND NOVEL

; TITLE OF INVENTION: VARIANTS OF ALTERNATIVE SPLICING

; FILE REFERENCE: 2786-0160P

; CURRENT APPLICATION NUMBER: US/09/778,927A

; CURRENT FILING DATE: 2001-02-08

; PRIOR APPLICATION NUMBER: IL 134453

; PRIOR FILING DATE: 2000-02-09

; PRIOR APPLICATION NUMBER: IL135341

; PRIOR FILING DATE: 2000-03-29

; NUMBER OF SEQ ID NOS: 81

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 53

; LENGTH: 254

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc.feature

; LOCATION: (1)-(254)

; OTHER INFORMATION: Xaa = any amino acid, unknown, or other

US-09-778-927A-53

Query Match

61.5%; Score 32; DB 10; Length 254;

Best Local Similarity 60.0%; Pred. No. 1.1e+02;

Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10

||| | | |

Db 74 EEVFPPLAMNY 83

## RESULT 14

US-10-024-066-2

; Sequence 2, Application US/10024066

; Patent No. US20020166134A1

; GENERAL INFORMATION:

; APPLICANT: Field, Loren J.

; APPLICANT: Pasumarthi, Kishore Babu S.

; TITLE OF INVENTION: CARDIOMYOCYTES WITH ENHANCED PROLIFERATIVE POTENTIAL,

; TITLE OF INVENTION: AND METHODS FOR PREPARING AND USING SAME

; FILE REFERENCE: 7037-450

; CURRENT APPLICATION NUMBER: US/10/024,066

; CURRENT FILING DATE: 2001-12-18

; PRIOR APPLICATION NUMBER: 60/139,942

; PRIOR FILING DATE: 1999-06-18

; PRIOR APPLICATION NUMBER: PCT/US00/16827

; PRIOR FILING DATE: 2000-06-19

; NUMBER OF SEQ ID NOS: 8

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 2

; LENGTH: 289

; TYPE: PRT

; ORGANISM: Mus musculus

US-10-024-066-2

Query Match

61.5%; Score 32; DB 9; Length 289;

Best Local Similarity 60.0%; Pred. No. 1.3e+02;

Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10

||| | | |

Db 73 EEVFPPLAMNY 82

## RESULT 15

US-10-024-066-4

; Sequence 4, Application US/10024066

; Patent No. US20020166134A1

; GENERAL INFORMATION:

; APPLICANT: Field, Loren J.

; APPLICANT: Pasumarthi, Kishore Babu S.

; TITLE OF INVENTION: CARDIOMYOCYTES WITH ENHANCED PROLIFERATIVE POTENTIAL,

; TITLE OF INVENTION: AND METHODS FOR PREPARING AND USING SAME

; FILE REFERENCE: 7037-450

; CURRENT APPLICATION NUMBER: US/10/024,066

; CURRENT FILING DATE: 2001-12-18

; PRIOR APPLICATION NUMBER: 60/139,942

; PRIOR FILING DATE: 1999-06-18

; PRIOR APPLICATION NUMBER: PCT/US00/16827

; PRIOR FILING DATE: 2000-06-19

; NUMBER OF SEQ ID NOS: 8

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 4

; LENGTH: 289

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-024-066-4

Query Match

61.5%; Score 32; DB 9; Length 289;

Best Local Similarity 60.0%; Pred. No. 1.3e+02;

Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10

||| | | |

Db 74 EEVFPPLAMNY 83

## RESULT 16

US-09-919-497-54

; Sequence 54, Application US/09919497

; Patent No. US20020108662A1

; GENERAL INFORMATION:

; APPLICANT: Mutter, George L.

; TITLE OF INVENTION: PROGNOSTIC CLASSIFICATION OF ENDOMETRIAL CANCER

; FILE REFERENCE: B0801/7225

; CURRENT APPLICATION NUMBER: US/09/919,497

; CURRENT FILING DATE: 2001-07-31

; PRIOR APPLICATION NUMBER: US 60/221,735

; PRIOR FILING DATE: 2000-07-31

; NUMBER OF SEQ ID NOS: 100

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 54

; LENGTH: 289

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-919-497-54

Query Match

61.5%; Score 32; DB 10; Length 289;

Best Local Similarity 60.0%; Pred. No. 1.3e+02;

Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10

||| | | |

Db 74 EEVFPPLAMNY 83

## RESULT 17

US-09-925-300-1061

; Sequence 1061, Application US/09925300

; Patent No. US20020151681A1

; GENERAL INFORMATION:

; APPLICANT: Craig Rosen,

; APPLICANT: Steve Ruben

; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies

; FILE REFERENCE: PA101

; CURRENT APPLICATION NUMBER: US/09/925,300

; CURRENT FILING DATE: 2001-08-10

; PRIOR APPLICATION NUMBER: PCT/US00/05988

; PRIOR FILING DATE: 2000-03-08

```

; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1061
; LENGTH: 295
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; LOCATION: (243)
; OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; NAME/KEY: SITE
; OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids
US-09-925-300-1061

Query Match          61.5%; Score 32; DB 10; Length 295;
Best Local Similarity 60.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      2 EVVPXGMSYS 11
Db      52 EVLPKMSYA 61

RESULT 18
US-09-923-304-4
; Sequence 4, Application US/09923304
; Patent No. US20020081612A1
; GENERAL INFORMATION:
; APPLICANT: KATZ, RUTH
; TITLE OF INVENTION: DETECTION AND DIAGNOSIS OF SMOKING RELATED CANCERS
; FILE REFERENCE: UTSC:658US
; CURRENT APPLICATION NUMBER: US/09/923,304
; CURRENT FILING DATE: 2001-08-06
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 529
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-923-304-4

Query Match          61.5%; Score 32; DB 10; Length 529;
Best Local Similarity 50.0%; Pred. No. 2.4e+02;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY      1 EEVVPXGMSY 10
Db      48 QEVLPPLGLKY 57

RESULT 19
US-10-101-921-4
; Sequence 4, Application US/10101921
; Publication No. US20030022199A1
; GENERAL INFORMATION:
; APPLICANT: Nezu, Jun-Ichi
; APPLICANT: Ose, Asuka
; APPLICANT: Tsuji, Akira
; TITLE OF INVENTION: TRANSPORTER GENES OATP-B, C, D, AND E
; FILE REFERENCE: 06501-104U51
; CURRENT APPLICATION NUMBER: US/10/101,921
; CURRENT FILING DATE: 2002-06-28
; PRIOR APPLICATION NUMBER: PCT/JP00/06416
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: JP 11/267835
; PRIOR FILING DATE: 1999-09-21
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4

```

```

; LENGTH: 691
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-101-921-4

Query Match          61.5%; Score 32; DB 9; Length 691;
Best Local Similarity 62.5%; Pred. No. 3.2e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      3 VVPXGMSY 10
Db      188 IVPLGLSY 195

RESULT 20
US-09-925-731-2
; Sequence 2, Application US/09925731
; Patent No. US20020090622A1
; GENERAL INFORMATION:
; APPLICANT: ADEKUN, ANTHONI MONISOLA
; APPLICANT: AMBROSE, HELEN JEAN
; APPLICANT: CRESSWELL, CARL JOHN
; APPLICANT: DUDLEY, ADAM JESTON
; TITLE OF INVENTION: CHEMICAL COMPOUNDS
; FILE REFERENCE: DJB/009901/0282795
; CURRENT APPLICATION NUMBER: US/09/925,731
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/226,909
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 691
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-731-2

Query Match          61.5%; Score 32; DB 10; Length 691;
Best Local Similarity 62.5%; Pred. No. 3.2e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      3 VVPXGMSY 10
Db      188 IVPLGLSY 195

RESULT 21
US-09-815-242-10384
; Sequence 10384, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625

```

; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; NUMBER OF SEQ ID NOS: 14110  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 10384  
; LENGTH: 1377  
; TYPE: PRT  
; ORGANISM: Escherichia coli  
US-09-815-242-10384

Query Match 61.5%; Score 32; DB 10; Length 1377;  
Best Local Similarity 45.5%; Pred. No. 6.7e+02;  
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEWXPXGMSYS 11  
|::| |::|  
Db 369 EQLNPAGLSYT 379

RESULT 22  
US-10-151-736-4  
; Sequence 4, Application US/10151736  
; Publication No. US20020192160A1  
; GENERAL INFORMATION:  
; APPLICANT: Callaghan, Michelle J.  
; APPLICANT: Sutherland, Lindfield  
; APPLICANT: Watts, Colin K.  
; TITLE OF INVENTION: No. US20020192160A1 Human Tumour Suppressor Gene  
; FILE REFERENCE: RICH-010CON  
; CURRENT APPLICATION NUMBER: US/10/151,736  
; PRIOR FILING DATE: 2002-05-15  
; PRIOR APPLICATION NUMBER: 09/403,402  
; PRIOR FILING DATE: 1999-10-21  
; PRIOR APPLICATION NUMBER: PCT/AU98/00280  
; PRIOR FILING DATE: 1998-04-20  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 2799  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-151-736-4

Query Match 61.5%; Score 32; DB 9; Length 2799;  
Best Local Similarity 60.9%; Pred. No. 1.4e+03;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 EVVXPXGMSYS 11  
||:| |::|  
Db 2096 EVLPTRKMSYA 2105

RESULT 23  
US-10-092-154-878  
; Sequence 878, Application US/10092154  
; Publication No. US20030054375A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PC009C1  
; CURRENT APPLICATION NUMBER: US/10/092,154  
; CURRENT FILING DATE: 2002-03-07  
; NUMBER OF SEQ ID NOS: 2003  
; PRIOR APPLICATION NUMBER: US/09/948,080  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 878  
; LENGTH: 53  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:

; NAME/KEY: misc\_feature  
; LOCATION: (7)  
; OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids  
US-10-092-154-878

Query Match 59.6%; Score 31; DB 9; Length 53;  
Best Local Similarity 66.7%; Pred. No. 33;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 VVPXGMSYS 11  
||| |::|  
Db 39 VVPTAVSYS 47

RESULT 24  
US-09-764-847-878  
; Sequence 878, Application US/09764847  
; Patent No. US20020132767A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PC009  
; CURRENT APPLICATION NUMBER: US/09/764,847  
; CURRENT FILING DATE: 2001-01-17  
; Prior application data removed - consult PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 2003  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 878  
; LENGTH: 53  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (7)  
; OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids  
US-09-764-847-878

Query Match 59.6%; Score 31; DB 10; Length 53;  
Best Local Similarity 66.7%; Pred. No. 33;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 VVPXGMSYS 11  
||| |::|  
Db 39 VVPTAVSYS 47

RESULT 25  
US-09-948-080-14  
; Sequence 14, Application US/09948080  
; Patent No. US20020102702A1  
; GENERAL INFORMATION:  
; APPLICANT: VAN DER OSTEN, CLAUS  
; APPLICANT: HALKIER, TORDEN  
; APPLICANT: ANDERSEN, CARSTEN  
; APPLICANT: BAUDITZ, PETER  
; APPLICANT: HANSEN, PETER KAMP  
; TITLE OF INVENTION: PROTEASE VARIANTS AND COMPOSITIONS  
; FILE REFERENCE: 4946,200-US  
; CURRENT APPLICATION NUMBER: US/09/948,080  
; CURRENT FILING DATE: 2001-09-06  
; PRIOR APPLICATION NUMBER: US/08/963,851  
; PRIOR FILING DATE: 1997-11-04  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 14  
; LENGTH: 59  
; TYPE: PRT  
; ORGANISM: Enterococcus faecalis  
US-09-948-080-14

Query Match 59.6%; Score 31; DB 10; Length 59;  
Best Local Similarity 45.5%; Pred. No. 37;  
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 BEVVPXGMSYS 11  
| : : | : |  
Db 38 EKHIPGGLEYS 48

Search completed: June 10, 2003, 14:35:40  
Job time : 16.0714 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2003, 13:31:15 ; Search time 11.2143 Seconds  
(without alignments)  
94.297 Million cell updates/sec

Title: US-09-909-164-5  
Perfect score: 52  
Sequence: 1 EEVVPXGMSYS 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5  
Searched: 283224 seqs, 96134422 residues  
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 75 summaries

Database : PIR\_73:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	38	73.1	3472	T31308	hypothetical 367K
2	37	71.2	840	T39116	probable sulfate p
3	37	71.2	877	T40413	sulfate permease -
4	36	69.2	102	A42452	V1 protein - tobac
5	36	69.2	1498	B97355	DNA segregation AT
6	35	67.3	225	S57810	hypothetical prote
7	35	67.3	425	T24111	hypothetical prote
8	35	67.3	670	S22293	zinc finger protei
9	35	67.3	749	H82691	topoisomerase IV s
10	35	67.3	2717	A34203	DNA-binding protei
11	34	65.4	156	S54619	hypothetical prote
12	34	65.4	252	H69491	cell division inhi
13	34	65.4	544	C82900	probable ABC subst
14	33	63.5	94	T40758	hypothetical prote
15	33	63.5	116	E90544	50S ribosomal prot
16	33	63.5	165	D69493	hypothetical prote
17	33	63.5	253	C81374	hypothetical prote
18	33	63.5	259	T34536	hypothetical prote
19	33	63.5	284	S75817	hypothetical prote
20	33	63.5	298	T47670	beta-ketoacyl-ACP
21	33	63.5	368	F72281	hypothetical prote
22	33	63.5	426	D82163	3-phosphoshikimate
23	33	63.5	466	T43653	cdc37 protein - fi
24	33	63.5	653	D82352	iron(III) ABC tran
25	33	63.5	890	A30481	bacteriocin BCN5 -
26	33	63.5	1028	AF3286	ATP-dependent DNA
27	33	63.5	1152	D87046	conserved hypothet
28	33	63.5	1394	A35626	transforming growt
29	33	63.5	1401	G82336	DNA-directed RNA p

30	63.5	1548	2	T04456	hypothetical prote	
31	63.5	1712	2	A38261	masking protein pr	
32	61.5	84	2	E97333	hypothetical prote	
33	61.5	175	2	P00616	transport protein	
34	61.5	223	2	T01457	rho protein GDP-di	
35	61.5	279	2	B72481	hypothetical prote	
36	61.5	288	2	JC4011	cyclin D2 - rat	
37	61.5	288	2	I58372	cyclin D2 - rat	
38	61.5	289	2	A41984	cyclin D2 - mouse	
39	61.5	289	2	A42822	cyclin D2 - human	
40	61.5	291	2	S57922	cyclin D1 - Africa	
41	61.5	291	2	S57925	cyclin D2 - Africa	
42	61.5	291	2	JC4579	cyclin D2 - chicke	
43	61.5	291	2	S62730	cyclin D1 - zebra	
44	61.5	292	2	B42822	cyclin D3 - human	
45	61.5	295	2	A38977	cyclin D1 - human	
46	61.5	295	2	A36523	cyclin D1 - mouse	
47	61.5	295	2	JC2342	cyclin D1 - rat	
48	61.5	347	2	I55120	rhsF protein - Esc	
49	61.5	363	2	D69551	conserved hypothet	
50	61.5	427	2	F64064	tolB protein - Hae	
51	61.5	449	2	A99286	conserved hypothet	
52	61.5	498	2	B90604	conserved hypothet	
53	61.5	525	2	D98311	agaA protein [mpo	
54	61.5	525	2	AF2971	hypothetical prote	
55	61.5	726	2	T44000	transport protein	
56	61.5	726	2	T44187	infected cell prot	
57	61.5	759	2	S25330	SCR1 protein - yea	
58	61.5	889	2	S22659	hypothetical prote	
59	61.5	922	2	AG1827	maltooligosyltreh	
60	61.5	993	1	P1VXTA	RNA 1 protein - to	
61	61.5	1306	2	S22624	aggregation protei	
62	61.5	1377	2	C65159	rhsA protein precu	
63	61.5	1377	2	E86034	rhsA protein in rh	
64	61.5	1394	2	H91236	RhsH core protein	
65	61.5	1397	2	A85570	rhsC protein in rh	
66	61.5	1397	2	C64805	rhsC protein precu	
67	61.5	1399	2	A99720	RhsC core protein	
68	61.5	1409	2	P91187	rhsA core protein	
69	61.5	1411	2	E65145	rhsB protein precu	
70	61.5	1948	2	S00485	gene 11-1 protein	
71	59.6	124	1	VKLJSI	trans-regulatory s	
72	31	59.6	133	2	D71173	hypothetical prote
73	31	59.6	208	1	D70764	probable cobH - My
74	31	59.6	224	2	G70709	probable purO prot
75	31	59.6	224	2	F87186	phosphoribosylform

## ALIGNMENTS

## RESULT 1

T31308  
hypothetical 367K protein - Cenarchaeum symbiosum

C:Species: Cenarchaeum symbiosum

C:Date: 11-Jan-2000 #sequence\_revision 11-Jan-2000 #text\_change 18-Feb-2000

R:Accession: T31308

J. Schleper, C.; DeLong, E.F.; Preston, C.M.; Feldman, R.A.; Wu, K.Y.; Swanson, R.V.

A:Reference number: Z20994; MUID:98422450; PMID:9748430

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-3472 <SCH>

A:Cross-references: EMBL:AF083072; NID:g3599393; PID:g3599394; PIDN:AAC62699.1

C:Superfamily: Cenarchaeum symbiosum hypothetical 367K protein

Query Match 73.1%; Score 38; DB 2; Length 3472;

Best Local Similarity 54.5%; Pred. No. 59;

Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

1 EEVVPXGMSYS 11

Db 2294 EDVIPRGISFS 2304  
 [:::| |:::|  
 RESULT 2  
 T39116  
 probable sulfate permease - fission yeast (Schizosaccharomyces pombe)  
 C:Species: Schizosaccharomyces pombe  
 C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Dec-1999  
 C:Accession: T39116  
 R:Hunt, C.; Aves, S.; McDougall, R.C.; Rajandream, M.A.; Barrell, B.G.  
 submitted to the EMBL Data Library, November 1999  
 A:Reference number: Z21829  
 A:Accession: T39116  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-840 <HUN>  
 A:Cross-references: EMBL:AL132779; PIDN:CAB60015.1; GSPDB:GN00066; SPDB:SPAC869.05C  
 A:Experimental source: strain 972h-; cosmid c869  
 C:Genetics:  
 A:Gene: SPDB:SPAC869.05C  
 A:Map position: 1  
 Query Match 71.2%; Score 37; DB 2; Length 840;  
 Best Local Similarity 77.8%; Pred. No. 21;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 3 VVPXGMSYS 11  
 [:::| |:::|  
 Db 135 VVPQGSYS 143  
 RESULT 3  
 T40413  
 sulfate permease - fission yeast (Schizosaccharomyces pombe)  
 C:Species: Schizosaccharomyces pombe  
 C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Dec-1999  
 C:Accession: T40413  
 R:Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Jimenez Martinez, J.  
 submitted to the EMBL Data Library, August 1998  
 A:Reference number: Z21926  
 A:Accession: T40413  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-877 <LYN>  
 A:Cross-references: EMBL:AL031261; PIDN:CAA20298.1; GSPDB:GN00067; SPDB:SPBC3H7.02  
 A:Experimental source: strain 972h-; cosmid c3H7  
 C:Genetics:  
 A:Gene: SPDB:SPBC3H7.02  
 A:Map position: 2  
 Query Match 71.2%; Score 37; DB 2; Length 877;  
 Best Local Similarity 77.8%; Pred. No. 22;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 3 VVPXGMSYS 11  
 [:::| |:::|  
 Db 148 VVPQGSYS 156  
 RESULT 4  
 A42452  
 V1 protein - tobacco yellow dwarf virus (strain Australia)  
 C:Species: tobacco yellow dwarf virus  
 C:Date: 15-Jan-1993 #sequence\_revision 15-Jan-1993 #text\_change 08-Oct-1999  
 C:Accession: A42452  
 R:Morris, B.A.M.; Richardson, K.A.; Haley, A.; Zhan, X.; Thomas, J.E.  
 Virology 187, 633-642, 1992  
 A:Title: The nucleotide sequence of the infectious cloned DNA component of tobacco yellow  
 A:Reference number: A42452; MUID:92188538; PMID:1546458  
 A:Accession: A42452  
 A:Molecule type: DNA  
 A:Residues: 1-102 <MOR>

A:Cross-references: GB:M81103; NID:g335283; PIDN:AAA47947.1; PID:g335284

Query Match 69.2%; Score 36; DB 2; Length 102;  
 Best Local Similarity 60.0%; Pred. No. 3.5;  
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 EVVPXGMSYS 11  
 [:::| |:::|  
 Db 7 QVVPNGINSYS 16

## RESULT 5

B97355

DNA segregation ATPase, FtsK/SpoIIIE family, YUKA B. subtilis ortholog [imported]  
 C:Species: Clostridium acetobutylicum  
 C:Date: 14-Sep-2001 #sequence\_revision 14-Sep-2001 #text\_change 14-Sep-2001  
 C:Accession: B97355  
 R:Nolling, J.; Breton, G.; Omeichenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.;  
 J. Bacteriol. 183, 4823-4838, 2001  
 A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium  
 A:Reference number: A96900; MUID:21359325; PMID:21359325  
 A:Accession: B97355  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-1498 <KUR>  
 A:Cross-references: GB:AE001437; PIDN:AAK81629.1; PID:g15026814; GSPDB:GN00168  
 A:Experimental source: Clostridium acetobutylicum ATCC824  
 C:Genetics:  
 A:Gene: CAC3709

Query Match 69.2%; Score 36; DB 2; Length 1498;  
 Best Local Similarity 60.0%; Pred. No. 63;  
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10  
 [:::| |:::|  
 Db 1276 EQIPMGMSY 1285

## RESULT 6

S57810

hypothetical protein precursor (clone TPP11) - tomato  
 C:Species: Lycopersicon esculentum (tomato)  
 C:Date: 28-Oct-1995 #sequence\_revision 03-Nov-1995 #text\_change 21-Jul-2000  
 C:Accession: S57810  
 R:Milligan, S.B.; Gasser, C.S.  
 Plant Mol. Biol. 28, 691-711, 1995  
 A:Title: Nature and regulation of pistil-expressed genes in tomato.  
 A:Reference number: S57808; MUID:95375233; PMID:7647301  
 A:Accession: S57810  
 A:Status: preliminary; nucleic acid sequence not shown  
 A:Molecule type: mRNA  
 A:Residues: 1-225 <ML>  
 A:Cross-references: EMBL:U20592; NID:g924625; PIDN:AAA80497.1; PID:g924626  
 C:Superfamily: plant kunitz-type proteinase inhibitor

Query Match 67.3%; Score 35; DB 2; Length 225;  
 Best Local Similarity 54.5%; Pred. No. 13;  
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMSYS 11  
 [:::| |:::|  
 Db 32 DEVVPNGKTYA 42

## RESULT 7

T24111

hypothetical protein R10D12.10 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
 C:Accession: T24111  
 R:Percy, C.

submitted to the EMBL Data Library, October 1996

A:Reference number: Z19842  
 A:Accession: T24111  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-425 <WIL>  
 A:Cross-references: EMBL:Z81109; PIDN:CAB03241.1; GSPDB:GN00023; CESP:R10D12.10  
 A:Experimental source: clone R10D12  
 C:Genetics:  
 A:Gene: CESP:R10D12.10  
 A:Map position: 5  
 A:Introns: 23/3; 56/3; 113/3; 257/2

Query Match 67.3%; Score 35; DB 2; Length 425;  
 Best Local Similarity 50.0%; Pred. No. 26;  
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10  
 I::: I: I  
 Db 335 EQIVPGGLQY 344

# RESULT 8

S22293  
 zinc finger protein AT-BP2 - rat (fragment)  
 C:Species: Rattus norvegicus (Norway rat)  
 C:Date: 29-Jan-1998 #sequence\_revision 06-Feb-1998 #text\_change 20-Sep-1999  
 C:Accession: S22293; I78656  
 R:MitcheImore, C.; Traboni, C.; Cortese, R.  
 Nucleic Acids Res. 19, 141-147, 1991  
 A:Title: Isolation of two cDNAs encoding zinc finger proteins which bind to the alpha 1-  
 A:Reference number: I58280; MUID:91187610; PMID:1901405

A:Accession: S22293  
 A:Status: nucleic acid sequence not shown  
 A:Molecule type: mRNA  
 A:Residues: 1-670 <MIT>  
 A:Cross-references: EMBL:X54250; NID:g57519; PIDN:CAA38151.1; PID:g57520  
 A:Note: the authors did not translate the codon for residue 1  
 C:Superfamily: HIV-EP2 enhancer-binding protein  
 C:Keywords: DNA binding; transcription regulation; zinc finger

Query Match 67.3%; Score 35; DB 2; Length 670;  
 Best Local Similarity 66.7%; Pred. No. 43;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMSYS 11  
 I::: I::: I  
 Db 376 VVPAGLTYS 384

# RESULT 9

H82691  
 topoisomerase IV subunit xfl353 [imported] - Xylella fastidiosa (strain 9a5c)  
 C:Species: Xylella fastidiosa  
 C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Sep-2000  
 C:Accession: H82691  
 R:anonymous. The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen

Nature 406, 151-157, 2000  
 A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.  
 A:Reference number: A82515; MUID:20365717; PMID:10910347  
 A:Note: for a complete list of authors see reference number A59328 below  
 A:Accession: H82691  
 A:Status: preliminary  
 A:Molecule type: DNA

A:Residues: 1-749 <SIM>  
 A:Cross-references: GB:AE003967; GB:AE003849; NID:g9106347; PIDN:AAF84162.1; GSPDB:GN001  
 R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A  
 Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Canargo, L.E.A.; Carraro, D.M.; Carre, H  
 as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.  
 submitted to GenBank, June 2000  
 A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm  
 J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig

chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martin  
 A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C  
 , F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri,  
 Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; da Sil  
 A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Sil  
 M.; Tshukako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.I

A:Reference number: A59328  
 A:Contents: annotation  
 C:Genetics:  
 A:Gene: XFL353  
 C:Superfamily: DNA topoisomerase (ATP-hydrolyzing) chain A; phage T4 DNA topoisom

Query Match 67.3%; Score 35; DB 2; Length 749;  
 Best Local Similarity 77.8%; Pred. No. 48;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 EVVPXGMSY 10  
 I::: I::: I  
 Db 526 EVDPSGMSY 534

# RESULT 10

A34203  
 DNA-binding protein PRDII-BF1 - human  
 N:Alternate names: major histocompatibility complex enhancer-binding protein 1  
 C:Species: Homo sapiens (man)  
 C:Date: 22-Jun-1990 #sequence\_revision 22-Jun-1990 #text\_change 20-Sep-1999  
 C:Accession: A34203; A34779  
 R:Fan, C.M.; Maniatis, T.  
 Genes Dev. 4, 29-42, 1990  
 A:Title: A DNA-binding protein containing two widely separated zinc finger motifs  
 A:Reference number: A34203; MUID:90169514; PMID:2106471

A:Accession: A34203  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-2717 <FAN>  
 A:Cross-references: EMBL:X51435; NID:g38017; PIDN:CAA35798.1; PID:g38018  
 R:Baldwin Jr., A.S.; LeClair, K.P.; Singh, H.; Sharp, P.A.  
 Mol. Cell Biol. 10, 1406-1414, 1990  
 A:Title: A large protein containing zinc finger domains binds to related sequence  
 A:Reference number: A34779; MUID:90205817; PMID:2108316

A:Accession: A34779  
 A:Status: preliminary; nucleic acid sequence not shown  
 A:Molecule type: mRNA  
 A:Residues: 801-1072, 'N', 1074-1168, 'K', 1170-1225, 'V', 1227-1434, 'N', 1436-1607, 'I', 16  
 A:Cross-references: GB:N32019  
 C:Superfamily: HIV-EP2 enhancer-binding protein  
 C:Keywords: DNA binding; transcription regulation; zinc finger

Query Match 67.3%; Score 35; DB 2; Length 2717;  
 Best Local Similarity 66.7%; Pred. No. 1.9e+02;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMSYS 11  
 I::: I::: I  
 Db 2405 VVPAGLTYS 2413

# RESULT 11

S54619  
 hypothetical protein YOR013w - yeast (Saccharomyces cerevisiae)  
 N:Alternate names: hypothetical protein O2612; hypothetical protein YOL303.3  
 C:Species: Saccharomyces cerevisiae  
 C:Date: 08-Jul-1995 #sequence\_revision 01-Sep-1995 #text\_change 19-Apr-2002  
 C:Accession: S54619; S68879  
 R:de Haan, M.; Maarse, A.C.; Grivell, L.A.  
 submitted to the EMBL Data Library, May 1995

A:Reference number: S54619  
 A:Accession: S54619  
 A:Molecule type: DNA  
 A:Residues: 1-156 <DBH>  
 A:Cross-references: EMBL:X87331; NID:g1041652; PIDN:CAA60762.1; PID:g899123  
 R:de Haan, M.; Grivell, L.A.; Maarse, A.C.

submitted to the Protein Sequence Database, July 1996  
 A:Reference number: S66877

A:Accession: S66879

A:Molecule type: DNA

A:Residues: 1-156 <DEW>

A:Cross-references: EMBL:Z74920; NID:gl420109; PIDN:CAA99201.1; PID:gl420111; MIPS:YOR01

A:Experimental source: strain S288C

C:Genetics:

A:Cross-references: SGD:S0005539

A:Map position: 15R

C:Superfamily: hypothetical protein YOR013w

Query Match 65.4%; Score 34; DB 2; Length 156;  
 Best Local Similarity 66.7%; Pred. No. 14;

Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 EVVPXGMSY 10  
 ||| |||

Db 50 EVMPGLGMDY 58  
 ||| |||

#### RESULT 12

H69491

cell division inhibitor (minD-2) homolog - Archaeoglobus fulgidus

C:Species: Archaeoglobus fulgidus

C>Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 26-Aug-1999

C:Accession: H69491

R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson

; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.

Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.

Nature 390, 364-370, 1997

A:Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.

Smith, H.O.; Woese, C.R.; Venter, J.C.

A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo

A:Reference number: A69250; MUID:98049343; PMID:9389475

A:Accession: H69491

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-252 <KLE>

A:Cross-references: GB:AE000970; GB:AE000782; NID:g2689293; PIDN:AAB89318.1; PID:g264860

C:Superfamily: cell division inhibitor mind

Query Match 65.4%; Score 34; DB 2; Length 252;  
 Best Local Similarity 75.0%; Pred. No. 24;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 EVVPXGMS 9  
 ||| |||

Db 81 EVIPAGMS 88  
 ||| |||

#### RESULT 13

C82900

probable ABC substrate-binding protein, iron UU359 [imported] - Ureaplasma urealyticum

C:Species: Ureaplasma urealyticum

C>Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Sep-2000

C:Accession: C82900

R:Glass, J.I.; Leikowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Cassell, G.H.

submitted to GenBank, February 2000

A:Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a min

A:Reference number: A82870

A:Accession: C82900

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-544 <GLA>

A:Cross-references: GB:AE002133; GB:AF222894; NID:g6899339; PIDN:AAF30768.1; GSPDB:GN001

A:Experimental source: serovar 3; biovar 1

C:Genetics:

A:Gene: ABCsbp-5; UU359

A:Genetic code: SGC3

Query Match

Best Local Similarity 65.4%; Score 34; DB 2; Length 544;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10  
 ||||| |||

Db 135 EEVVPXGMSY 144  
 ||||| |||

#### RESULT 14

I40758

hypothetical protein 1 - Campylobacter jejuni (fragment)

C:Species: Campylobacter jejuni

C>Date: 16-Aug-1996 #sequence\_revision 16-Aug-1996 #text\_change 08-Oct-1999

C:Accession: I40758; S47317

R:Hani, E.K.; Chan, V.L.

J. Bacteriol. 177, 2396-2402, 1995

A:Title: Expression and characterization of Campylobacter jejuni benzoylglycine amide

A:Reference number: I40758; MUID:95247673; PMID:7730270

A:Accession: I40758

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-94 <RES>

A:Cross-references: EMBL:Z36940; NID:g535805; PIDN:CAA85392.1; PID:g535806

Query Match 63.5%; Score 33; DB 2; Length 94;

Best Local Similarity 55.6%; Pred. No. 13;

Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 EVVPXGMSY 10  
 :: |||||

Db 26 DIPPSGMSY 34

#### RESULT 15

E90544

50S ribosomal protein L20 [imported] - Mycoplasma pulmonis (strain UAB CTIP)

C:Species: Mycoplasma pulmonis

C>Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 03-Aug-2001

C:Accession: E90544

R:Chambaud, I.; Heilig, R.; Ferris, S.; Barbe, V.; Samson, D.; Galisson, F.; Moszer,

Nucleic Acids Res. 29, 2145-2153, 2001

A:Title: The complete genome sequence of the murine respiratory pathogen Mycoplasma

A:Reference number: A99512; MUID:21267165; PMID:11353084

A:Accession: E90544

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-116 <KUR>

A:Cross-references: GB:AL445566; PID:gl4089674; PIDN:CAC13434.1; GSPDB:GN00153

A:Experimental source: strain UAB CTIP

C:Genetics:

A:Gene: MYP0\_2610

A:Genetic code: SGC3

C:Superfamily: Escherichia coli ribosomal protein L20

Query Match 63.5%; Score 33; DB 2; Length 116;

Best Local Similarity 77.8%; Pred. No. 17;

Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 VVPXGMSYS 11  
 ||| |||||

Db 68 VRPLGMSYS 76

#### RESULT 16

D69493

hypothetical protein AF1949 - Archaeoglobus fulgidus

C:Species: Archaeoglobus fulgidus

C>Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 04-Mar-2000

C:Accession: D69493

R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dou

; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, M.

Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.

Nature 390, 364-370, 1997

A:Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes



Smith, H.O.; Woese, C.R.; Venter, J.C.  
A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeon  
A:Reference number: A69250; MUID:98049343; PMID:9389475  
A:Accession: D69493  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-165 <KLE>  
A:Cross-references: GB:AE000968; GB:AE000782; NID:g2689291; PIDN:AAB89307.1; PID:g264859  
C:Superfamily: Archaeoglobus fulgidus hypothetical protein AF1949

Query Match 63.5%; Score 33; DB 2; Length 165;  
Best Local Similarity 60.0%; Pred. No. 25;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EVVXPXGMSY 10  
II I I I I I  
Db 60 EESIPDGASY 69

RESULT 17  
C81374  
hypothetical protein Cj0990c [imported] - Campylobacter jejuni (strain NCTC 11168)  
C:Species: Campylobacter jejuni  
C:Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 03-Jun-2002  
C:Accession: C81374  
R:Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chilling  
C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Barrell  
Nature 403, 665-668, 2000  
A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hyp  
A:Reference number: A81250; MUID:20150912; PMID:10688204  
A:Accession: C81374  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-253 <PAR>  
A:Cross-references: GB:AL1139076; GB:AL111168; NID:g96968128; PIDN:CAB73246.1; PID:g9696842  
A:Experimental source: serotype O2, strain NCTC 11168  
C:Genetics:  
A:Gene: Cj0990c

Query Match 63.5%; Score 33; DB 2; Length 253;  
Best Local Similarity 55.6%; Pred. No. 39;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 EVVXPXGMSY 10  
:: I I I I I  
Db 185 DIFPSGMSY 193

RESULT 18  
T34536  
hypothetical protein DKF2p434C031.1 - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 29-Oct-1999  
C:Accession: T34536  
R:Poustka, A.; Wellenreuther, R.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.  
submitted to the Protein Sequence Database, October 1999  
A:Reference number: Z21540  
A:Accession: T34536  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-259 <POU>  
A:Cross-references: EMBL:AL122063  
A:Experimental source: adult testis; clone DKF2p434C031  
C:Genetics:  
A:Note: DKF2p434C031.1

Query Match 63.5%; Score 33; DB 2; Length 259;  
Best Local Similarity 60.0%; Pred. No. 40;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 EVVXPXGMSY 11  
II I I I I I  
Db 22 EVAPAGASYN 31

Smith, H.O.; Woese, C.R.; Venter, J.C.  
A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeon  
A:Reference number: A69250; MUID:98049343; PMID:9389475  
A:Accession: D69493  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-165 <KLE>  
A:Cross-references: GB:AE000968; GB:AE000782; NID:g2689291; PIDN:AAB89307.1; PID:g264859  
C:Superfamily: Archaeoglobus fulgidus hypothetical protein AF1949

Query Match 63.5%; Score 33; DB 2; Length 165;  
Best Local Similarity 60.0%; Pred. No. 25;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EVVXPXGMSY 10  
II I I I I I  
Db 60 EESIPDGASY 69

RESULT 17  
C81374  
hypothetical protein Cj0990c [imported] - Campylobacter jejuni (strain NCTC 11168)  
C:Species: Campylobacter jejuni  
C:Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 03-Jun-2002  
C:Accession: C81374  
R:Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chilling  
C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Barrell  
Nature 403, 665-668, 2000  
A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hyp  
A:Reference number: A81250; MUID:20150912; PMID:10688204  
A:Accession: C81374  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-253 <PAR>  
A:Cross-references: GB:AL1139076; GB:AL111168; NID:g96968128; PIDN:CAB73246.1; PID:g9696842  
A:Experimental source: serotype O2, strain NCTC 11168  
C:Genetics:  
A:Gene: Cj0990c

Query Match 63.5%; Score 33; DB 2; Length 253;  
Best Local Similarity 55.6%; Pred. No. 39;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 EVVXPXGMSY 10  
:: I I I I I  
Db 185 DIFPSGMSY 193

RESULT 18  
T34536  
hypothetical protein DKF2p434C031.1 - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 29-Oct-1999  
C:Accession: T34536  
R:Poustka, A.; Wellenreuther, R.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.  
submitted to the Protein Sequence Database, October 1999  
A:Reference number: Z21540  
A:Accession: T34536  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-259 <POU>  
A:Cross-references: EMBL:AL122063  
A:Experimental source: adult testis; clone DKF2p434C031  
C:Genetics:  
A:Note: DKF2p434C031.1

Query Match 63.5%; Score 33; DB 2; Length 259;  
Best Local Similarity 60.0%; Pred. No. 40;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 EVVXPXGMSY 11  
II I I I I I  
Db 22 EVAPAGASYN 31

## RESULT 19

S75817

hypothetical protein slr1275 - Synechocystis sp. (strain PCC 6803)

C:Species: Synechocystis sp.

A:Variety: PCC 6803

C:Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 08-Oct-1999

C:Accession: S75817

K:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima

O, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Y

DNA Res. 3, 109-136, 1996

A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechoc

S.

A:Reference number: S74322; MUID:97061201; PMID:8905231

A:Accession: S75817

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-284 &lt;KAN&gt;

A:Cross-references: EMBL:D90913; GB:AB001339; NID:g1653348; PIDN:BAA18276.1; PID:d1

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 63.5%; Score 33; DB 2; Length 284;

Best Local Similarity 55.6%; Pred. No. 44;

Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMSYS 11

I I I I I

Db 208 VIPAGVSYT 216

## RESULT 20

T47670

beta-ketoacyl-ACP reductase-like protein - Arabidopsis thaliana

N:Alternate names: protein T26112.190

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 18-Aug-2000

C:Accession: T47670

R:Monfort, A.; Casacuberta, E.; Puigdomenech, P.; Mewes, H.W.; Lemcke, K.; Mayer, K

submitted to the Protein Sequence Database, February 2000

A:Reference number: Z24471

A:Accession: T47670

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-298 &lt;MON&gt;

A:Cross-references: EMBL:AL132954

A:Experimental source: cultivar Columbia; BAC clone T26112

C:Genetics:

A:Map position: 3

A:Introns: 25/3

A:Note: T26112.190

C:Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology

Query Match

Best Local Similarity 55.6%; Score 33; DB 2; Length 298;

Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMSYS 11

I I I I I

Db 187 IVPGLAYS 195

## RESULT 21

F72281

hypothetical protein TM1216 - Thermotoga maritima (strain MSB8)

C:Species: Thermotoga maritima

C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 21-Jul-2000

C:Accession: F72281

R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; H

Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson

C.M.

Nature 399, 323-329, 1999

A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genom

A;Reference number: A72200; MUID:99287316; PMID:10360571  
 A;Accession: F72281  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-368 <AR>  
 A;Cross-references: GB:AE001778; GB:AE000512; NID:g4981757; PIDN:AAD36291.1; PID:g498176  
 A;Experimental source: strain MSB8  
 C;Genetics:  
 A;Gene: TM1216  
 C;Superfamily: NADH dehydrogenase (ubiquinone) 49K protein

Query Match 63.5%; Score 33; DB 2; Length 368;  
 Best Local Similarity 55.6%; Pred. No. 59;  
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMSYS 11  
 :|| ||:|  
 Db 294 IVPKGMAYA 302

## RESULT 22

D82163  
 C;Species: 3-phosphoshikimate 1-carboxyvinyltransferase VC1732 [Imported] - Vibrio cholerae (strain  
 C;Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Feb-2001  
 C;Accession: D82163  
 R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;  
 Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, H.  
 l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
 Nature 406, 477-483, 2000  
 A;Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.  
 A;Reference number: A82035; MUID:20406833; PMID:10952301

A;Accession: D82163  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-426 <HEI>  
 A;Cross-references: GB:AE004251; GB:AE003852; NID:g9656248; PIDN:AAF94882.1; GSPDB:GN001  
 A;Experimental source: serogroup O1; strain N16961; biotype El Tor  
 C;Genetics:  
 A;Gene: VC1732  
 A;Map position: 1  
 C;Superfamily: 3-phosphoshikimate 1-carboxyvinyltransferase; 3-phosphoshikimate 1-carboxy

Query Match 63.5%; Score 33; DB 2; Length 426;  
 Best Local Similarity 60.0%; Pred. No. 69;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10  
 :||:| ||  
 Db 223 EFVIPAGQSY 232

## RESULT 23

T43653  
 cdc37 protein - fission yeast (Schizosaccharomyces pombe)  
 C;Species: Schizosaccharomyces pombe  
 C;Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 18-Feb-2000  
 C;Accession: T43653; T40791; T43654  
 R;Westwood, P.K.; Preston, N.C.; Fantes, P.A.  
 submitted to the EMBL Data Library, March 1999  
 A;Description: Schizosaccharomyces pombe cdc37 cDNA.  
 A;Reference number: 222602

A;Accession: T43653  
 A;Status: preliminary; translated from GB/EMBL/DBJ  
 A;Molecule type: mRNA  
 A;Residues: 1-466 <WES>  
 A;Cross-references: EMBL:AJ132376; PIDN:CAB38757.1  
 R;Wood, V.; Rajandream, M.A.; Barrell, B.G.; Volckaert, G.  
 submitted to the EMBL Data Library, May 1999  
 A;Reference number: 221875  
 A;Accession: T40791  
 A;Status: preliminary; translated from GB/EMBL/DBJ  
 A;Molecule type: DNA

A;Residues: 1-466 <WOO>  
 A;Cross-references: EMBL:AL049769; PIDN:CAB42371.2; GSPDB:GN00067  
 A;Experimental source: strain 972h-; cosmid c9B6  
 R;Westwood, P.K.; Preston, N.C.; Fantes, P.A.  
 submitted to the EMBL Data Library, March 1999  
 A;Description: Schizosaccharomyces pombe cdc37 gene.  
 A;Reference number: 222603

A;Accession: T43654  
 A;Status: preliminary; translated from GB/EMBL/DBJ  
 A;Molecule type: DNA  
 A;Residues: 1-466 <WE2>  
 A;Cross-references: EMBL:AJ132377; PIDN:CAB38758.1  
 C;Genetics:  
 A;Gene: cdc37; SPAC9B6.10  
 A;Map position: 2  
 A;Introns: 8/2; 17/2; 21/1

Query Match 63.5%; Score 33; DB 2; Length 466;  
 Best Local Similarity 50.0%; Pred. No. 76;  
 Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10  
 :|| ||||  
 Db 98 DSAIPGMSY 107

## RESULT 24

D82352  
 C;Species: ABC transporter, permease protein VC0203 [Imported] - Vibrio cholerae (str  
 C;Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Feb-2001  
 C;Accession: D82352  
 R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.  
 chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Seller  
 l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
 Nature 406, 477-483, 2000  
 A;Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.  
 A;Reference number: A82035; MUID:20406833; PMID:10952301

A;Accession: D82352  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-653 <HEI>  
 A;Cross-references: GB:AE004110; GB:AE003852; NID:g9654600; PIDN:AAF93379.1; GSPDB:G  
 A;Experimental source: serogroup O1; strain N16961; biotype El Tor  
 C;Genetics:  
 A;Gene: VC0203  
 A;Map position: 1

Query Match 63.5%; Score 33; DB 2; Length 653;  
 Best Local Similarity 66.7%; Pred. No. 11e+02;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXGMS 9  
 :|| ||||:|  
 Db 300 EEVVPSGIT 308

## RESULT 25

A30481  
 bacteriocin BCN5 - Clostridium perfringens plasmid pIP404  
 C;Species: Clostridium perfringens  
 C;Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 24-Nov-1999  
 C;Accession: A30481; S03779  
 R;Garnier, T.; Cole, S.T.  
 J. Bacteriol. 168, 1189-1196, 1986  
 A;Title: Characterization of a bacteriocinogenic plasmid from Clostridium perfringens

A;Reference number: JT0354; MUID:87057020; PMID:2877971  
 A;Accession: A30481  
 A;Molecule type: DNA  
 A;Residues: 1-890 <GAR>  
 A;Cross-references: GB:M32882; GB:J03309; NID:g150738; PIDN:AAA98249.1; PID:g150739  
 C;Genetics:  
 A;Gene: bcn

A;Genome: plasmid  
C;Superfamily: Clostridium perfringens plasmid pIP404 bacteriocin BCN5  
C;Keywords: bacteriocin

Query Match 63.5%; Score 33; DB 2; Length 890;  
Best Local Similarity 66.7%; Pred. NO. 1.5e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 EVVPXGMSY 10  
|||||  
Db 170 EVVPGGFTY 178

Search completed: June 10, 2003, 13:49:09  
Job time : 13.2143 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2003, 13:25:04 ; Search time 4.5 Seconds  
(without alignments)  
101.387 Million cell updates/sec

Title: US-09-909-164-5  
Perfect score: 52  
Sequence: 1 EGVVPXGMSYS 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 75 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	38	73.1	1058	1 CARB_FUSNN	Q8R986 fusbacteri
2	37	71.2	877	1 SULH_SCHPO	O74377 schizosacch
3	36	69.2	102	1 Y11K_TVDVA	P31619 tobacco yel
4	36	69.2	1498	1 Y1A9_GLOAB	O04351 clostridium
5	35	67.3	2717	1 ZEP1_HUMAN	P15822 homo sapien
6	34	65.4	788	1 CY14_NEUCR	P23622 neurospora
7	34	65.4	1499	1 A10C_HUMAN	O60312 homo sapien
8	33	63.5	116	1 RL20_MYCPU	Q98qvo mycoplasma
9	33	63.5	253	1 Y990_CAMJE	P45489 campylobact
10	33	63.5	426	1 AROA_VIBCH	Q9Ktb0 vibrio chol
11	33	63.5	466	1 CC37_SCHPO	O94740 schizosacch
12	33	63.5	478	1 GSR2_HUMAN	Q9nm5 homo sapien
13	33	63.5	890	1 BCN5_CLOPE	P08696 clostridium
14	33	63.5	1394	1 LTBS_HUMAN	P22064 homo sapien
15	33	63.5	1401	1 RPOC_VIBCH	Q9Kv29 vibrio chol
16	33	63.5	1595	1 LTBL_HUMAN	O14766 homo sapien
17	33	63.5	1712	1 LTB1_RAT	Q00918 rattus norv
18	32	61.5	288	1 CGD2_RAT	Q04827 rattus norv
19	32	61.5	289	1 CGD2_HUMAN	P30279 homo sapien
20	32	61.5	289	1 CGD2_MOUSE	P30280 mus musculu
21	32	61.5	291	1 CGD1_BRARE	Q90459 brachydanio
22	32	61.5	291	1 CGD1_XENLA	P50755 xenopus lae
23	32	61.5	291	1 CGD2_CHICK	P49706 gallus gall
24	32	61.5	291	1 CGD2_XENLA	P53782 xenopus lae
25	32	61.5	292	1 CGD1_CHICK	P55169 gallus gall
26	32	61.5	292	1 CGD3_HUMAN	P30281 homo sapien
27	32	61.5	295	1 CGD1_HUMAN	P24385 homo sapien
28	32	61.5	295	1 CGD1_MOUSE	P25322 mus musculu
29	32	61.5	295	1 CGD1_RAT	P39948 rattus norv
30	32	61.5	427	1 TOLB_HAEIN	P44677 haemophilus
31	32	61.5	529	1 ENP3_HUMAN	O75355 homo sapien
32	32	61.5	691	1 OAT6_HUMAN	Q9Y616 homo sapien
33	32	61.5	726	1 PRTP_HSV6U	P52384 human herpe

34	61.5	759	1 SCTL_YEAST	P32784 saccharomyc
35	61.5	920	1 EDD_RAT	Q62671 rattus norv
36	61.5	993	1 V1A_TAV	P28931 tomato aspe
37	61.5	1377	1 RHA_ECOLI	P16916 escherichia
38	61.5	1397	1 RHSC_ECOLI	P16918 escherichia
39	61.5	1411	1 RHSE_ECOLI	P16917 escherichia
40	61.5	2799	1 EDD_HUMAN	O95071 homo sapien
41	59.6	124	1 REV_SIVCZ	P17280 chimpanzee
42	59.6	208	1 COBH_MYCTU	O10676 mycobacteri
43	59.6	223	1 PURO_PYRHO	O59619 pyrococcus
44	59.6	224	1 PURO_HALNI	Q9Hn2 halobacteri
45	59.6	224	1 PURO_MYCLE	O05756 mycobacteri
46	59.6	224	1 PURO_MYCTU	P71841 mycobacteri
47	59.6	225	1 PURO_CORAM	Q9rhx0 corynebacte
48	59.6	240	1 GDJR_ARATH	Q9sf66 arabidopsis
49	59.6	276	1 Y939_METJA	O58349 methanococc
50	59.6	319	1 YHAI_CRYPA	P10941 cryptoneutr
51	59.6	432	1 PURA_YEAST	P80210 saccharomyc
52	59.6	488	1 NOM2_PSEAE	Q9htr0 pseudomonas
53	59.6	670	1 OATP_RAT	P46720 rattus norv
54	59.6	706	1 ADDG_HUMAN	Q9uey8 homo sapien
55	59.6	706	1 ADDG_MOUSE	Q9qyb5 mus musculu
56	59.6	827	1 PLSE_MOUSE	Q61586 mus musculu
57	59.6	828	1 PLSE_RAT	P97564 rattus norv
58	59.6	1047	1 EF3_SCHPO	O94489 schizosacch
59	59.6	1407	1 RPOC_ECOLI	P00577 escherichia
60	59.6	1426	1 RHSD_ECOLI	P16919 escherichia
61	59.6	2145	1 U520_CAEEL	Q9u240 caenorhabdl
62	58.7	472	1 ET2A_XENLA	P19102 xenopus lae
63	57.7	81	1 YE47_ARCFU	O28825 archaeoglob
64	57.7	121	1 TKNK_HUMAN	Q9uhf0 homo sapien
65	57.7	132	1 ATPE_AQUAE	O66903 aquifex aeo
66	57.7	146	1 ATPE_LACAC	Q9rgv0 lactobacill
67	57.7	218	1 PURO_METH	O26270 methanobact
68	57.7	223	1 PURO_RHIME	Q92p11 rhizobium m
69	57.7	223	1 RPIA_BUCAI	P57489 buchnera ap
70	57.7	230	1 PURO_METJA	O59042 methanococc
71	57.7	232	1 SCOA_HELPJ	O921e3 helicobacte
72	57.7	232	1 SCOA_HELPJ	P56006 helicobacte
73	57.7	286	1 CXA6_RAT	P28233 rattus norv
74	57.7	356	1 GBA2_USMTA	P87033 utillago ma
75	57.7	361	1 RFBB_SALTY	P26391 salmonella

## ALIGNMENTS

RESULT 1  
CARB\_FUSNN STANDARD; PRT; 1058 AA.  
ID CARB\_FUSNN STANDARD; PRT; 1058 AA.  
AC Q8R986;  
DT 15-JUN-2002 (Rel. 41, Created)  
DT 15-JUN-2002 (Rel. 41, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Carbamoyl-phosphate synthase large chain (EC 6.3.5.5) (Carbamoyl-phosphate synthase ammonia chain).  
GN CARB OR FN0422.  
OS Fusobacterium nucleatum (subsp. nucleatum).  
OC Bacteria; Fusobacteriia; Fusobacterium.  
OX NCBI\_TaxID=76856;  
RN [1]  
RC SEQUENCE FROM N.A.  
RX MEDLINE=21886394; PubMed=11889109;  
RA Kapatal V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A., Bhattacharyya A., Bartman A., Gardner W., Grechkin G., Zhu L., Vasieva O., Chu L., Kogan Y., Chaga O., Goltsman E., Bernal A., Larsen N., D'Souza M., Walunas T., Pusch G., Haselkorn R., Fonstein M., Kyrpides N., Overbeek R.;  
RT "Genome sequence and analysis of the oral bacterium Fusobacterium nucleatum strain ATCC 25586.";  
RL J. Bacteriol. 184:2003-2018(2002).  
CC -I- CATALYTIC ACTIVITY: 2 ATP + L-glutamine + CO(2) + H(2)O = 2 ADP +

phosphate + L-glutamate + carbamoyl phosphate.  
 -!- COFACTOR: Binds three manganese ions (By similarity).  
 -!- PATHWAY: Arginine biosynthesis.  
 -!- SUBUNIT: Composed of two chains; the small (or glutamine) chain promotes the hydrolysis of glutamine to ammonia, which is used by the large (or ammonia) chain to synthesize carbamoyl phosphate (By similarity).  
 -!- SIMILARITY: BELONGS TO THE CARB FAMILY.  
 -----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 EMBL; AE010554; AAL94625.1; ALT\_INIT.  
 InterPro; IPR005483; CPase\_L.  
 InterPro; IPR005479; CPase\_L\_D2.  
 InterPro; IPR005480; CPase\_L\_D3.  
 InterPro; IPR005481; CPase\_L\_N.  
 InterPro; IPR004362; MGS\_like.  
 Pfam; PF02789; CPase\_L\_chain; 2.  
 Pfam; PF02786; CPase\_L\_D2; 2.  
 Pfam; PF02787; CPase\_L\_D3; 1.  
 Pfam; PF02142; MGS; 1.  
 PRINTS; PR00098; CPASE.  
 PROSITE; PS00866; CPASE\_1; 2.  
 PROSITE; PS00867; CPASE\_2; 2.  
 Arginine biosynthesis; Pyrimidine biosynthesis; Ligase; Repeat;  
 ATP-binding; Manganese; Complete proteome.  
 FT DOMAIN 1 401  
 FT DOMAIN 402 546  
 FT DOMAIN 547 929  
 FT DOMAIN 930 1058  
 FT REPEAT 1 546  
 FT REPEAT 547 1058  
 FT NP\_BIND 153 210  
 FT NP\_BIND 302 352  
 FT NP\_BIND 384 284  
 FT METAL 298 298  
 FT METAL 300 300  
 FT METAL 820 820  
 FT METAL 832 832  
 FT METAL 832 832  
 SQ SEQUENCE 1058 AA; 117451 MW; ED7037AF77C1E39F CRC64;  
 -----  
 Query Match 73.1%; Score 38; DB 1; Length 1058;  
 Best Local Similarity 60.0%; Pred. No. 6.2;  
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
 QY 2 EVVPGXGMSYS 11  
 Db 190 EIVPGLNYS 199  
 |::|::|  
 RESULT 2  
 SULH\_SCHPO  
 ID SULH\_SCHPO STANDARD; PRT; 877 AA.  
 AC 074377;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Probable sulfate permease C3H7.02.  
 GN SPBC3H7.02.  
 OS Schizosaccharomyces pombe (Fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OX NCBI\_TaxID=4896;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RC STRAIN-972;  
 RX MEDLINE=21848401; PubMed=11859360;  
 RA WOOD V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,  
 RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,  
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,  
 RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,  
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,  
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,  
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,  
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell J.,  
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,  
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,  
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,  
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
 RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,  
 RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,  
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,  
 RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,  
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,  
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,  
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
 RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,  
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
 RA Cerretti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,  
 RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;  
 RT "The genome sequence of Schizosaccharomyces pombe.";  
 RL Nature 415:871-880(2002).  
 CC -!- FUNCTION: HIGH AFFINITY UPTAKE OF SULFATE INTO THE CELL (BY  
 CC SIMILARITY).  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
 CC -!- SIMILARITY: BELONGS TO THE SLC26A FAMILY OF TRANSPORTERS.  
 CC -!- SIMILARITY: CONTAINS 1 STAS DOMAIN.  
 CC -----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 EMBL; AL031261; CAA20298.1;  
 CC InterPro; IPR002645; STAS.  
 DR InterPro; IPR001902; Sulfate\_transp.  
 DR Pfam; PF00916; Sulfate\_transp; 1.  
 DR Pfam; PF01740; STAS; 1.  
 DR TIGRFAMs; TIGR00815; sulp; 1.  
 DR PROSITE; PS01130; SLC26A; 1.  
 DR PROSITE; PS0801; STAS; 1.  
 KW Transport; Transmembrane.  
 FT TRANSMEM 133 153  
 FT TRANSMEM 161 181  
 FT TRANSMEM 186 206  
 FT TRANSMEM 221 241  
 FT TRANSMEM 243 263  
 FT TRANSMEM 292 312  
 FT TRANSMEM 329 349  
 FT TRANSMEM 384 404  
 FT TRANSMEM 424 444  
 FT TRANSMEM 461 481  
 FT TRANSMEM 484 504  
 FT TRANSMEM 518 538  
 FT TRANSMEM 543 563  
 FT DOMAIN 594 747  
 STAS.  
 SQ SEQUENCE 877 AA; 96373 MW; 56995A8493371E43 CRC64;  
 -----  
 Query Match 71.2%; Score 37; DB 1; Length 877;  
 Best Local Similarity 77.8%; Pred. No. 8.3;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 3 VVPXGMSYS 11  
 ||| |||||

Db 148 VVPGMSYA 156

## RESULT 3

Y1LK\_TYDVA STANDARD; PRT; 102 AA.

AC P31619;

DT 01-JUL-1993 (Rel. 26, Created)

DT 01-JUL-1993 (Rel. 26, Last sequence update)

DT 01-OCT-1993 (Rel. 27, Last annotation update)

DE Hypothetical 11.2 kDa protein.

GN V1.

OS Tobacco yellow dwarf virus (strain Australia) (TYDV).

OC Viruses; ssDNA viruses; Geminiviridae; Mastrevirus.

OX NCBI\_TaxID=31599;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=92188538; PubMed=1546458;

RA Morris B.A.M., Richardson K.A., Haley A., Zhan X., Thomas J.E.;

RT "The nucleotide sequence of the infectious cloned DNA component of

RT tobacco yellow dwarf virus reveals features of geminiviruses

RT infecting monocotyledonous plants.";

RL Virology 187:633-642(1992).

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; M81103; AAA47947.1; -.

DR PIR; A42452; A42452.

DR InterPro; IPR002621; Gemini\_mov.

DR Pfam; PF01708; Gemini\_mov; 1.

KW Hypothetical protein.

SQ SEQUENCE 102 AA; 11178 MW; A40ECF1E0AF55B67 CRC64;

Query Match

Best Local Similarity 69.2%; Score 36; DB 1; Length 102;

Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 EVVPXGMSYS 11

DB 7 QVVPSCINYS 16

## RESULT 4

Y1A9\_CLOB

ID Y1A9\_CLOB STANDARD; PRT; 1498 AA.

AC Q04351;

DT 01-FEB-1994 (Rel. 28, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Hypothetical protein CAC3709.

GN CAC3709

OS Clostridium acetobutylicum.

OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;

OC Clostridium.

OX NCBI\_TaxID=1488;

RN [1]

RP SEQUENCE FROM N.A.

RX STRAIN=ATCC 824 / DSM 792 / VKM B-1787;

RA MEDLINE=21359325; PubMed=11466286;

RA Neelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q.,

RA Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I.,

RA Tatunov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,

RA Bennett G.N., Koonin E.V., Smith D.R.;

RT "Genome sequence and comparative analysis of the solvent-producing

RT bacterium Clostridium acetobutylicum.";

RL J. Bacteriol. 183:4823-4838(2001).

RN [2]

RP SEQUENCE OF 1-108 FROM N.A.

RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;

RX MEDLINE=93273706; PubMed=8501044;

RA Sauer U., Duerre P.;

RT "Sequence and molecular characterization of a DNA region encoding a

RT small heat shock protein of Clostridium acetobutylicum.";

RL J. Bacteriol. 175:3394-3400(1993).

CC -!- SIMILARITY: BELONGS TO THE FTSK/SPOIITE FAMILY.

CC -!- CAUTION: Ref.2 sequence differs from that shown due to frameshifts

CC in positions 76 and 106.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; AE007866; AAK81629.1; -.

DR EMBL; X65276; CA446379.1; ALT\_FRAME.

DR InterPro; IPR002543; FtsK\_SpoIIIE.

DR Pfam; PF01580; FtsK\_SpoIIIE; 2.

KW Hypothetical protein; ATP-binding; Complete proteome.

FT NP\_BIND 675 682 ATP (POTENTIAL).

SQ SEQUENCE 1498 AA; 168968 MW; FF42037A335A9649 CRC64;

Query Match

Best Local Similarity 69.2%; Score 36; DB 1; Length 1498;

Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10

DB 1276 EQKIPGMSY 1285

## RESULT 5

ZEPI\_HUMAN

ID ZEPI\_HUMAN STANDARD; PRT; 2717 AA.

AC P15822;

DT 01-APR-1990 (Rel. 14, Created)

DT 01-APR-1990 (Rel. 14, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Zinc finger protein 40 (Human immunodeficiency virus type I enhancer-

DE binding protein 1) (HIV-Ep1) (Major histocompatibility complex binding

DE protein 1) (MBP-1) (Positive regulatory domain II binding factor 1)

DE (PRDII-BF1).

GN HIVEP1 OR ZNF40.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=90169514; PubMed=2106471;

RA Fan C.M., Maniatis T.;

RT "A DNA-binding protein containing two widely separated zinc finger

RT motifs that recognize the same DNA sequence.";

RL Genes Dev. 4:29-42(1990).

RN [2]

RP STRUCTURE BY NMR OF 2113-2142.

RX MEDLINE=91064333; PubMed=2248949;

RA Omichinski J.G., Clore G.M., Appella E., Sakaguchi K.,

RA Gronenborn A.M.;

RT "High-resolution three-dimensional structure of a single zinc finger

RT from a human enhancer binding protein in solution.";

RL Biochemistry 29:9324-9334(1990).

RN [3]

RP STRUCTURE BY NMR OF 2087-2142.

RX MEDLINE=92232684; PubMed=1567844;

RA Omichinski J.G., Clore G.M., Robien M., Sakaguchi K., Appella E.,

RA Gronenborn A.M.;

RT "High-resolution solution structure of the double Cys2His2 zinc

finger from the human enhancer binding protein MBP-1.";  
 RT Biochemistry 31:3907-3917(1992).  
 CC -1- FUNCTION: THIS PROTEIN SPECIFICALLY BINDS TO THE DNA SEQUENCE  
 CC 5'-GGGACTTCC-3' WHICH IS FOUND IN THE ENHANCER ELEMENTS OF  
 CC NUMEROUS VIRAL PROMOTERS SUCH AS THOSE OF SV40, CMV, OR HIV1.  
 CC IN ADDITION, RELATED SEQUENCES ARE FOUND IN THE ENHANCER ELEMENTS  
 CC OF A NUMBER OF CELLULAR PROMOTERS, INCLUDING THOSE OF THE CLASS I  
 CC MHC, INTERLEUKIN-2 RECEPTOR, AND INTERFERON-BETA GENES. IT MAY ACT  
 CC IN T-CELL ACTIVATION.  
 CC -1- SUBCELLULAR LOCATION: Nuclear.  
 CC -1- INDUCTION: BY MITOGEN AND PHORBOL ESTER.  
 CC -1- DOMAIN: CONTAINS TWO SETS OF 2 ZINC-FINGERS, WHICH ARE WIDELY  
 CC SEPARATED AND RECOGNIZE THE SAME DNA SEQUENCE. THERE IS A FIFTH  
 CC ZINC-FINGER IN-BETWEEN.  
 CC -1- SIMILARITY: STRONG, TO HIVEP2.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; X51435; CAA35798.1; -  
 CC PIR; A34203; A34203.  
 CC PDB; 3ZNF; 15-JAN-92.  
 CC PDB; 4ZNF; 15-JAN-92.  
 CC PDB; 1BBO; 31-OCT-93.  
 CC TRANSFAC; T00497; -  
 CC Genew; HGNC:4920; HIVEP1.  
 CC MIN; 194540; -  
 CC InterPro: IPR000822; Znf\_C2H2.  
 CC Pfam; PF00096; Zf\_C2H2; 5.  
 CC PRINTS; PR00048; ZNCFINGER.  
 CC SMART; SM00355; Znf\_C2H2; 4.  
 CC PROSITE; PS00028; ZINC\_FINGER\_C2H2\_1; 4.  
 CC PROSITE; PS0157; ZINC\_FINGER\_C2H2\_2; 4.  
 CC Transcription regulation; Zinc-finger; Metal-binding; DNA-binding;  
 CC Nuclear protein; Repeat; 3D-structure.  
 CC DOMAIN 406 456  
 CC ZN\_FING 406 428  
 CC ZN\_FING 434 456  
 CC C2H2-TYPE.  
 CC DOMAIN 803 806  
 CC POLX-SER.  
 CC ZN\_FING 958 981  
 CC C2HC-TYPE (POTENTIAL).  
 CC DOMAIN 2087 2139  
 CC ZN\_FING 2087 2109  
 CC ZN\_FING 2115 2139  
 CC ZN\_FING 2115 2139  
 CC TURN 2088 2088  
 CC TURN 2090 2092  
 CC STRAND 2095 2095  
 CC STRAND 2099 2108  
 CC HELIX 2109 2109  
 CC STRAND 2115 2116  
 CC STRAND 2123 2124  
 CC HELIX 2127 2135  
 CC SEQUENCE 2717 AA; D45D3CA951FEA561 CRC64;  
 SQ  
 Query Match 67.3%; Score 35; DB 1; Length 2717;  
 Best Local Similarity 66.7%; Pred. No. 70;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 QY 3 VVPXGMSYS 11  
 Db 2405 VVPAGLTYS 2413  
 RESULT 6  
 CY14\_NEUCR  
 ID CY14\_NEUCR STANDARD; PRT; 788 AA.  
 AC P23622;  
 DT 01-NOV-1991 (Rel. 20, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)

15-JUN-2002 (Rel. 41, Last annotation update)  
 DT Sulfate permease II.  
 DE GN Cys-14.  
 OS Neurospora crassa.  
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
 OC Sordariales; Sordariaceae; Neurospora.  
 OX NCBI\_TaxID=5141;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91129256; PubMed=1825178;  
 RA Ketter J.S., Jaraal G., Fu Y.-H., Marzluf G.A.;  
 RT "Nucleotide sequence, messenger RNA stability, and DNA recognition  
 RT elements of cys-14, the structural gene for sulfate permease II in  
 RT Neurospora crassa.";  
 RL Biochemistry 30:1780-1787(1991).  
 RN [2]  
 RP PROBABLE REVISIONS.  
 RX MEDLINE=94188926; PubMed=8140616;  
 RA Sandal N.N., Marcker K.A.;  
 RT "Similarities between a soybean nodulin, Neurospora crassa sulphate  
 RT permease II and a putative human tumour suppressor.";  
 RL Trends Biochem. Sci. 19:19-19(1994).  
 CC -1- FUNCTION: UPTAKE OF SULFATE INTO THE CELL.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -1- INDUCTION: Highly expressed, but only in cells subject to sulfur  
 CC limitation, and it is turned on by the positive-acting Cys-3  
 CC sulfur regulatory protein.  
 CC -1- MISCELLANEOUS: SULFATE PERMEASE II IS MAINLY FOUND IN MYCELIA.  
 CC -1- SIMILARITY: BELONGS TO THE SLC26A FAMILY OF TRANSPORTERS.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; M59167; AAA33615.1; ALT\_SEQ.  
 CC PIR; A37956; A37956.  
 CC InterPro: IPR001902; Sulfate transp.  
 CC Pfam; PF00916; Sulfate.transp; 1.  
 CC TIGRfams; TIGR00815; sulp; 1.  
 CC PROSITE; PS01130; SLC26A; 1.  
 CC Transport; Transmembrane; Glycoprotein.  
 CC TRANSMEM 71 91  
 CC TRANSMEM 103 123  
 CC TRANSMEM 128 148  
 CC TRANSMEM 171 191  
 CC TRANSMEM 193 213  
 CC TRANSMEM 271 291  
 CC TRANSMEM 326 346  
 CC TRANSMEM 363 383  
 CC TRANSMEM 451 471  
 CC TRANSMEM 474 494  
 CC CARBOHYD 23 23  
 CC CARBOHYD 578 578  
 CC SEQUENCE 788 AA; 87864 MW; 4FC604B60798CE77 CRC64;  
 SQ  
 Query Match 65.4%; Score 34; DB 1; Length 788;  
 Best Local Similarity 66.7%; Pred. No. 32;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 QY 3 VVPXGMSYS 11  
 Db 90 VVPOGMAYA 98  
 RESULT 7  
 A10C\_HUMAN  
 ID A10C\_HUMAN STANDARD; PRT; 1499 AA.  
 AC O60312; Q96914;  
 DT 30-MAY-2000 (Rel. 39, Created)



DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DE 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Potential phospholipid-transferring ATPase VC (EC 3.6.3.1) (ATPVC)  
 GN ATP10C OR ATPVC OR KIAA0566.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=2125279; PubMed=11326269;  
 RA Meguro M., Kashiwagi A., Mitsuya K., Nakao M., Kondo I., Saitoh S.,  
 RA Oshimura M.;  
 RT "A novel maternally expressed gene, ATP10C, encodes a putative  
 RT aminophospholipid translocase associated with Angelman syndrome.";  
 RL Nat. Genet. 28:19-20(2001).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=2131119; PubMed=11353404;  
 RA Herzog L.B.K., Kim S.-J., Cook E.H. Jr., Ledbetter D.H.;  
 RT "The human aminophospholipid-transferring ATPase gene ATP10C maps  
 RT adjacent to UBE3A and exhibits similar imprinted expression.";  
 RL Am. J. Hum. Genet. 68:1501-1505(2001).  
 RN [3]  
 RP SEQUENCE OF 337-1499 FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=98290545; PubMed=9628581;  
 RA Nagase T., Ishikawa K.-I., Miyajima N., Tanaka A., Kotani H.,  
 RA Nomura N., Ohara O.;  
 RT "Prediction of the coding sequences of unidentified human genes. IX.  
 RT The complete sequences of 100 new cDNA clones from brain which can  
 RT code for large proteins in vitro.";  
 RL DNA Res. 5:31-39(1998).  
 CC -!- CATALYTIC ACTIVITY: ATP + H(2)O -> ADP + phosphate.  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).  
 CC -!- TISSUE SPECIFICITY: Wide expression, with highest levels in  
 CC kidney, followed by lung, brain, prostate, testis, ovary, and  
 CC small intestine.  
 CC -!- DISEASE: Defects in ATP10C may be a cause of Angelman syndrome  
 CC (AS), also known as 'happy puppet syndrome'.  
 CC -!- SIMILARITY: BELONGS TO THE CATION TRANSPORT ATPASES FAMILY (E1-E2  
 CC ATPASES). SUBFAMILY IV.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AB051358; BAB47392.1; -;  
 DR EMBL: AY029504; AAK33100.1; -;  
 DR EMBL: AY029487; AAK33100.1; JOINED.  
 DR EMBL: AY029488; AAK33100.1; JOINED.  
 DR EMBL: AY029489; AAK33100.1; JOINED.  
 DR EMBL: AY029490; AAK33100.1; JOINED.  
 DR EMBL: AY029491; AAK33100.1; JOINED.  
 DR EMBL: AY029492; AAK33100.1; JOINED.  
 DR EMBL: AY029493; AAK33100.1; JOINED.  
 DR EMBL: AY029494; AAK33100.1; JOINED.  
 DR EMBL: AY029495; AAK33100.1; JOINED.  
 DR EMBL: AY029496; AAK33100.1; JOINED.  
 DR EMBL: AY029497; AAK33100.1; JOINED.  
 DR EMBL: AY029498; AAK33100.1; JOINED.  
 DR EMBL: AY029499; AAK33100.1; JOINED.  
 DR EMBL: AY029500; AAK33100.1; JOINED.  
 DR EMBL: AY029501; AAK33100.1; JOINED.  
 DR EMBL: AY029502; AAK33100.1; JOINED.  
 DR EMBL: AY029503; AAK33100.1; JOINED.  
 DR EMBL: AB011138; BAA25492.1; -;  
 DR Genew: HGNC:13547; ATP10C.

DR MIM: 605855; -;  
 DR MIM: 105830; -;  
 DR InterPro: IPR001757; ATPase\_E1-E2.  
 DR InterPro: IPR001454; Hlgase/hydrase.  
 DR Pfam: PF00702; Hydrolase; 1.  
 DR PRINTS: PR00119; CATATPASE.  
 DR PROSITE: PS00154; ATPASE\_E1\_E2; 1.  
 KW Hydrolase; Transmembrane; Phosphorylation; Magnesium; ATP-binding;  
 KW Multigene family.  
 FT DOMAIN 1 86 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 87 106 POTENTIAL.  
 FT DOMAIN 107 110 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 111 128 POTENTIAL.  
 FT DOMAIN 129 309 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 310 332 POTENTIAL.  
 FT DOMAIN 337 362 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 363 384 POTENTIAL.  
 FT DOMAIN 385 1087 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 1088 1108 POTENTIAL.  
 FT DOMAIN 1109 1119 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 1120 1140 POTENTIAL.  
 FT DOMAIN 1141 1170 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 1171 1192 POTENTIAL.  
 FT DOMAIN 1193 1199 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 1200 1222 POTENTIAL.  
 FT DOMAIN 1223 1228 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 1229 1249 POTENTIAL.  
 FT DOMAIN 1250 1267 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 1268 1292 POTENTIAL.  
 FT DOMAIN 1293 1499 CYTOPLASMIC (POTENTIAL).  
 FT MOD\_RES 427 427 PHOSPHORYLATION (BY SIMILARITY).  
 FT METAL 1031 1031 MAGNESIUM (BY SIMILARITY).  
 FT METAL 1035 1035 MAGNESIUM (BY SIMILARITY).  
 FT DOMAIN 467 470 POLY-GLU.  
 FT CONFLICT 388 388 Q -> R (IN REF. 3).  
 SQ SEQUENCE 1499 AA; 167687 MW; D4996A4D0635A68D CRC64;  
 Query Match 65.4%; Score 34; DB 1; Length 1499;  
 Best Local Similarity 72.7%; Pred. No. 62;  
 Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 OY 1 EEVVPXGMSYS 11  
 Db 469 EEVVPXGMSYS 479  
 |||||  
 RESULT 8  
 RL20\_MYCPU  
 ID RL20\_MYCPU STANDARD; PRT; 116 AA.  
 AC Q98QV0;  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE 50S ribosomal protein L20.  
 GN RPLT OR MYPU\_2610.  
 OS Mycoplasma pulmonis.  
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.  
 OX NCBI\_TaxID=2107;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=UAB CTIP;  
 RX MEDLINE=21267165; PubMed=11353084;  
 RA Chambaud I., Heilig R., Ferris S., Barbe V., Samson D., Galisson F.,  
 RA Moszer I., Dybvig K., Wroblewski H., Viari A., Rocha E.P.C.,  
 RA Blanchard A.;  
 RT "The complete genome sequence of the murine respiratory pathogen  
 RT Mycoplasma pulmonis.";  
 RL Nucleic Acids Res. 29:2145-2153(2001).  
 CC -!- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 23S RIBOSOMAL RNA AND IS  
 CC NECESSARY TO THE IN VITRO ASSEMBLY PROCESS OF THE 50S RIBOSOMAL  
 CC SUBUNIT; IT IS NOT INVOLVED IN THE PROTEIN SYNTHESIZING FUNCTION  
 CC OF THAT SUBUNIT (BY SIMILARITY).  
 CC -!- SIMILARITY: BELONGS TO THE L20P FAMILY OF RIBOSOMAL PROTEINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; AL445563; CAC13434.1; -  
 DR MyDulList; MYPU 2610; -  
 DR InterPro: IPR001081; Ribosomal\_L20.  
 DR Pfam; PF00453; Ribosomal\_L20; 1.  
 DR PRINTS; PR00062; Ribosomal\_L20.  
 DR ProDom; PD002389; Ribosomal\_L20; 1.  
 DR TIGRFAMs; TIGR01032; rplT\_bact; 1.  
 DR PROSITE; PS00937; RIBOSOMAL\_L20; 1.  
 DR Ribosomal protein; rRNA-binding; Complete proteome.  
 KW Ribosomal protein; rRNA-binding; Complete proteome.  
 SQ SEQUENCE 116 AA; 13565 MW; C59C748901B18F14 CRC64;

Query Match 63.5%; Score 33; DB 1; Length 116;  
 Best Local Similarity 77.8%; Pred. No. 7;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 VVPXGMSYS 11  
 I I I I I I I  
 DB 68 VRPLGMSYS 76

RESULT 9  
 Y990\_CAMJE STANDARD; PRT; 253 AA.  
 ID Y990\_CAMJE  
 AC P45489; Q9PNV0;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Hypothetical protein Cj0990c.  
 GN Cj0990c.  
 OS Campylobacter jejuni.  
 OC Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group;  
 OC Campylobacter.  
 OX NCBI\_TaxID=197;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NCTC 11168;  
 RX MEDLINE=20150912; PubMed=1068204;  
 RA Parkhill J., Wren B.W., Mungall K., Ketley J.M., Churcher C.,  
 RA Basham D., Chillingworth T., Davies R.M., Feltwell T., Holtrold S.,  
 RA Jagels K., Karlyshev A.V., Moule S., Pallen M.J., Penn C.W.,  
 RA Quail M.A., Rajandream M.A., Rutherford K.M., van Vliet A.H.M.,  
 RA Whitehead S., Barrell B.G.;  
 RT "The genome sequence of the food-borne pathogen Campylobacter jejuni  
 RT reveals hypervariable sequences.";  
 RL Nature 403:665-668(2000).  
 RN [2]  
 RP SEQUENCE OF 160-253 FROM N.A.  
 RC STRAIN=ATCC 43431 / TGH 9011;  
 RX MEDLINE=95247673; PubMed=7730270;  
 RA Hani E.K., Chan V.L.;  
 RT "Expression and characterization of Campylobacter jejuni  
 RT benzoylglycine amidohydrolase (Hippuricase) gene in Escherichia  
 RT coli.";  
 RL J. Bacteriol. 177:2396-2402(1995).  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; AL139076; CAB73246.1; -

DR EMBL; Z36940; CAA85392.1; -  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 253 AA; 29783 MW; F96D3FF3265F8A6A CRC64;

Query Match 63.5%; Score 33; DB 1; Length 253;  
 Best Local Similarity 55.6%; Pred. No. 16;  
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 EVVPXGMSY 10  
 I I I I I I I  
 DB 185 DIFPSGMSY 193

RESULT 10  
 AROA\_VIBCH STANDARD; PRT; 426 AA.  
 ID AROA\_VIBCH  
 AC C9KRE0;  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE 3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) (5-  
 DE enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) (EPSPS).  
 GN AROA OR VC1732.  
 OS Vibrio cholerae.  
 OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.  
 OX NCBI\_TaxID=666;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=El Tor N16961 / Serotype O1;  
 RX MEDLINE=20406833; PubMed=10952301;  
 RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,  
 RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,  
 RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,  
 RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,  
 RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,  
 RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,  
 RA Fraser C.M.;  
 RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio  
 RT cholerae.";  
 RL Nature 406:477-483(2000).  
 CC -1- CATALYTIC ACTIVITY: Phosphoenolpyruvate + 3-phosphoshikimate =  
 CC phosphate + 5-O-(1-carboxyvinyl)-3-phosphoshikimate.  
 CC -1- PATHWAY: Aromatic amino acids biosynthesis; shikimate pathway;  
 CC sixth step.  
 CC -1- SUBUNIT: Monomer (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).  
 CC -1- SIMILARITY: BELONGS TO THE EPSP SYNTHASE FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; AE004251; AAF94882.1; -  
 DR TIGR; VC1732; -  
 DR InterPro: IPR001986; EPSP\_synthase.  
 DR Pfam; PF002075; EPSP\_synthase; 1.  
 DR ProDom; PD001867; EPSP\_synthase; 1.  
 DR PROSITE; PS00104; EPSP\_SYNTHASE\_1; 1.  
 DR PROSITE; PS00885; EPSP\_SYNTHASE\_2; 1.  
 KW Aromatic amino acid biosynthesis; Transferase; Complete proteome.  
 SQ SEQUENCE 426 AA; 46101 MW; 38852D6483BFE1C3 CRC64;

Query Match 63.5%; Score 33; DB 1; Length 426;  
 Best Local Similarity 60.08%; Pred. No. 27;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10  
 I I I I I I I  
 DB 223 EFVIPAGQSY 232

Query Match 63.5%; Score 33; DB 1; Length 466;  
 Best Local Similarity 50.0%; Pred. No. 30;  
 Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10  
 : : : : :  
 Db 98 DSAIPGMSY 107

## RESULT 12

GSR2\_HUMAN  
 ID GSR2\_HUMAN STANDARD; PRT: 478 AA.  
 AC Q9NZM5; Q9NP1; Q9NP4; Q9UFI2; Q9BTC6; Q9HAX6;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Glioma tumor suppressor candidate region gene 2 protein (p60).  
 GN GLTSCR2.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.  
 RX MEDLINE=20175430; PubMed=10708517;  
 RA Smith J.S., Tachibana I., Pohl U., Lee H.K., Thanarajasingam U.,  
 RA Portier B.P., Ueki K., Billings S., Ramaswamy S., Mohrenweiser H.W.,  
 RA Scheithauer B.W., Louis D.N., Jenkins R.B.;  
 RT "A transcript map of the chromosome 19q-Arm glioma tumor suppressor  
 RT region.";  
 RL Genomics 64:44-50(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Muscle;  
 RA Strausberg R.;  
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE OF 9-478 FROM N.A., AND SUBCELLULAR LOCATION.  
 RX MEDLINE=99214318; PubMed=10196275;  
 RA Bruni R., Fineschi B., Ogle W.O., Roizman B.;  
 RT "A novel cellular protein, p60, interacting with both herpes simplex  
 RT virus 1 regulatory proteins ICP22 and ICP0 is modified in a  
 RT cell-type-specific manner and is recruited to the nucleus after  
 RT infection.";  
 RL J. Virol. 73:3810-3817(1999).  
 RN [4]  
 RP SEQUENCE OF 12-478 FROM N.A.  
 RA Andreu N., Estivill X., Escarceller M., Sumoy L.;  
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
 RN [5]  
 RP SEQUENCE OF 218-477 FROM N.A.  
 RC TISSUE=Testis;  
 RA Poustka A., Klein M., Mewes H.-W., Gassenhuber J., Wiemann S.;  
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.  
 CC -!- SUBUNIT: Interacts with HSV-1 early proteins ICP22 and ICP0.  
 CC -!- SUBCELLULAR LOCATION: Nuclear.  
 CC -!- TISSUE SPECIFICITY: Expressed at high levels in heart and  
 CC pancreas, moderate levels in placenta, liver, skeletal muscle, and  
 CC kidney, and low levels in brain and lung.  
 CC -!- SIMILARITY: BELONGS TO THE GLTSCR2 FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; AJ132377; CAB38758.1; -;  
 EMBL; AJ132376; CAB38757.1; -;  
 EMBL; AL049769; CAB42371.2; -;  
 KW Chapterone; Cell division; Cell cycle.  
 SQ SEQUENCE 466 AA; 52554 MW; 647238B34CABB3C5 CRC64;

RESULT 11  
 CC37\_SCHPO  
 ID CC37\_SCHPO STANDARD; PRT: 466 AA.  
 AC O94740;  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Hsp90 co-chaperone Cdc37 (Hsp90 chaperone protein kinase-targeting  
 DE subunit) (Cell division control protein 37).  
 GN CDC37 OR SPAG9B6.10.  
 OS Schizosaccharomyces pombe (Fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OC Schizosaccharomyces.  
 OX NCBI\_TaxID=4896;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Westwood P.K., Preston N.C., Fantes P.A.;  
 RT "Schizosaccharomyces pombe cdc37 gene.";  
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=972;  
 RX MEDLINE=21848401; PubMed=11859360;  
 RA Wood V., Williams R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,  
 RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,  
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,  
 RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,  
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,  
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,  
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,  
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,  
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,  
 RA Rutherford J., Rutter S., Saunders D., Seeger K., Sharp S.,  
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,  
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
 RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,  
 RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,  
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,  
 RA Borzym K., Langer I., Beck A., Leherach H., Reinhardt R., Pohl T.M.,  
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,  
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,  
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
 RA Lucas M., Rochet M., Gallard C., Tallada V.A., Garzon A., Thode G.,  
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
 RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,  
 RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;  
 RT "The genome sequence of Schizosaccharomyces pombe.";  
 RL Nature 415:871-880(2002).  
 CC -!- FUNCTION: With Hsp90 it forms a complex that binds to several  
 CC kinases, resulting in stabilization and promotion of their  
 CC activity (By similarity).  
 CC -!- SUBUNIT: Forms a complex with Hsp90. Interacts with a number of  
 CC kinases (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -!- SIMILARITY: BELONGS TO THE CDC37 FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; AJ132377; CAB38758.1; -;  
 EMBL; AJ132376; CAB38757.1; -;  
 EMBL; AL049769; CAB42371.2; -;  
 KW Chapterone; Cell division; Cell cycle.  
 SQ SEQUENCE 466 AA; 52554 MW; 647238B34CABB3C5 CRC64;

```

DR EMBL; BC010095; AAH10095.1; -
DR EMBL; AF296124; AAG30413.1; -
DR EMBL; AL359335; CAB94786.1; -
DR EMBL; AL359336; CAB94787.1; -
DR EMBL; AL122063; CAB59242.1; -
DR SWISS-2DPAGE; Q9NZM5; HUMAN.
DR Genew; HGNC:4333; GLUTSCR2.
DR MIM; 605691; -
KW Nuclear protein; Polymorphism.
FT VARIANT 389 389
R -> Q.
/FTID-VAR_011486.
FT CONFLICT 4 6
FT CONFLICT 9 9
FT CONFLICT 146 191
R -> R (IN REF. 3).
REKEOLWEKLAQSGELPREYRAQARLLNPSATRAKPGQD
TVERP -> SGRSSTGRSWPSRASSPGGAQSPVAQPFEN
KGNPAPAGHRIAA (IN REF. 3).
SDNPLDRPLVGQDEFFLE -> LNNPDKPVWPGCLFPG
(IN REF. 3).
A -> S (IN REF. 2; AAH04229).
D -> H (IN REF. 3).
PEGNLRDRFKSFQRNMIEPRERAKFRKRYKVLVEKRAF
REIQ -> VLVTSVCRGAPCVMTPTSLPPVPPRGYGRHHGCP
WAGPVGPMPRG (IN REF. 5).
EGNLRDRFKSFQRNMIEPRERAKFRKRYKVLVEKRAF
EIQL -> RGHSFETGSRAPRGGI (IN REF. 3).
SQ SEQUENCE 478 AA; 54417 MW; 7F18923E348CB52B CRC64;

Query Match 63.5%; Score 33; DB 1; Length 478;
Best Local Similarity 60.0%; Pred. No. 30;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 EVVPXGMSYS 11
||| | |
DB 239 EVAPAGASYN 248

RESULT 13
BCN5_CLOPE
ID BCN5_CLOPE STANDARD; PRT; 890 AA.
AC P08696;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last annotation update)
DE Bacteriocin BCN5.
GN BCN.
OS Clostridium perfringens.
OG Plasmid pip404.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1502;
RP [1]
RN SEQUENCE FROM N.A.
RC STRAIN=CPN50;
RX MEDLINE=88336297; PubMed=2901768;
RA Garnier T., Cole S.T.;
RT "Complete nucleotide sequence and genetic organization of the
RL bacteriocinogenic plasmid, pip404, from Clostridium perfringens.";
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CPN50;
RX MEDLINE=87057020; PubMed=2877971;
RA Garnier T., Cole S.T.;
RT "Characterization of a bacteriocinogenic plasmid from Clostridium
RL perfringens and molecular genetic analysis of the
RN bacteriocin-encoding gene.";
RP J. Bacteriol. 168:1189-1196(1986).
RN [3]
RP SEQUENCE OF 1-14 FROM N.A.
RC STRAIN=CPN50;
RX MEDLINE=89039249; PubMed=2460717;
RA Garnier T., Cole S.T.;
RT "Studies of UV-inducible promoters from Clostridium perfringens in

```

```

RT vivo and in vitro.";
RL Mol. Microbiol. 2:607-614(1988).
CC -!- FUNCTION: MAY FUNCTION AS AN IONOPHORE.
CC -!- INDUCTION: BY UV IRRADIATION.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M14481; AAA98248.1; -
DR EMBL; M32882; AAA98249.1; -
DR PIR; A30481; A30481.
DR InterPro; IPR003646; SH3_bac.
DR SMART; SM00287; SH3b; 3. Plasmid.
KW Antibiotic; Bacteriocin; Plasmid.
FT DOMAIN 815 869 HYDROPHOBIC.
SQ SEQUENCE 890 AA; 96699 MW; F4E5B8971C31C6C6 CRC64;

Query Match 63.5%; Score 33; DB 1; Length 890;
Best Local Similarity 66.7%; Pred. No. 58;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 EVVPXGMSY 10
|||| |
DB 170 EVVPGGFTY 178

RESULT 14
LTBS_HUMAN
ID LTBS_HUMAN STANDARD; PRT; 1394 AA.
AC P22064;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Latent transforming growth factor beta binding protein 15 precursor
DE (Transforming growth factor beta-1 binding protein 1) (TGF-beta1-BP-
DE 1).
GN LTBP1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RP [1]
RN SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Fibroblast, and Platelet;
RX MEDLINE=90275601; PubMed=2350783;
RA Kanzaki T., Olofsson A., Moren A., Wernstedt C., Hellman U.,
RA Miyazono K., Claesson-Welsh L., Heidin C.-H.;
RT "TGF-beta 1 binding protein: a component of the large latent complex
of TGF-beta 1 with multiple repeat sequences.";
RL Cell 61:1051-1061(1990).
CC -!- SUBUNIT: THE LARGE LATENT COMPLEX OF TGF-BETA1 FROM PLATELETS IS
CC COMPOSED OF THE TGF-BETA1 MOLECULE NONCOVALENTLY ASSOCIATED WITH
CC A DISULFIDE-BONDED COMPLEX OF A DIMER OF THE N-TERMINAL PROPEPTIDE
CC OF THE TGF-BETA1 PRECURSOR AND A THIRD COMPONENT DENOTED TGF-
CC BETA1-BP. TGF-BETA1-BP DOES NOT BIND DIRECTLY TO ACTIVE TGF-BETA1.
CC -!- ALTERNATIVE PRODUCTS: 2 isoforms; a short form (shown here) and a
CC long form (AC Q14766); are produced by alternative splicing.
CC -!- PTM: CONTAINS HYDROXYLATED ASPARAGINE RESIDUES.
CC -!- PTM: The N-terminus is blocked.
CC -!- SIMILARITY: CONTAINS 16 EGF-LIKE DOMAINS.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).

```

```

CC EMBL; M34057; AAA61160.1; -.
DR PIR; A35626; A35626.
DR HSP; P00750; LTGC.
DR GlycoSuiteDB; P22064; -.
DR Genew; HGNC:6714; LTBP1.
DR MIM; 150390; -.
DR InterPro; IPR000152; Asx_hydroxyl.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR002212; Fibril-assoc.
DR Pfam; PF00008; EGF; 15.
DR Pfam; PF00683; TB; 4.
DR SMART; SM00179; EGF_CA; 13.
DR PROSITE; PS00001; EGF_like; 4.
DR PROSITE; PS00010; ASX_HYDROXYL; 13.
DR PROSITE; PS00022; EGF_1; 2.
DR PROSITE; PS01186; EGF_2; 11.
DR PROSITE; PS01187; EGF_CA; 15.
KW Growth factor binding; Repeat; EGF-like domain; Hydroxylation; Signal;
KW Glycoprotein; Alternative splicing.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 1394 LATENT TRANSFORMING GROWTH FACTOR BETA
FT DOMAIN 300 340 EGF-LIKE 1, CALCIUM-BINDING (POTENTIAL).
FT REPEAT 348 412 REPEAT A.
FT DOMAIN 546 587 EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 588 629 EGF-LIKE 3, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 630 670 EGF-LIKE 4, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 671 710 EGF-LIKE 5, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 711 751 EGF-LIKE 6, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 752 792 EGF-LIKE 7, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 793 833 EGF-LIKE 8, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 834 874 EGF-LIKE 9, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 875 916 EGF-LIKE 10, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 917 958 EGF-LIKE 11, CALCIUM-BINDING (POTENTIAL).
FT REPEAT 1017 1084 EGF-LIKE 12, CALCIUM-BINDING (POTENTIAL).
FT REPEAT 1097 1139 REPEAT B.
FT REPEAT 1190 1262 REPEAT C.
FT DOMAIN 1294 1334 EGF-LIKE 14, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 1335 1379 EGF-LIKE 16, CALCIUM-BINDING (POTENTIAL).
FT SITE 847 849 CELL ATTACHMENT SITE (POTENTIAL).
FT DISULFID 304 315 BY SIMILARITY.
FT DISULFID 310 324 BY SIMILARITY.
FT DISULFID 326 339 BY SIMILARITY.
FT DISULFID 550 562 BY SIMILARITY.
FT DISULFID 557 571 BY SIMILARITY.
FT DISULFID 573 586 BY SIMILARITY.
FT DISULFID 592 604 BY SIMILARITY.
FT DISULFID 599 613 BY SIMILARITY.
FT DISULFID 615 628 BY SIMILARITY.
FT DISULFID 634 645 BY SIMILARITY.
FT DISULFID 640 654 BY SIMILARITY.
FT DISULFID 657 669 BY SIMILARITY.
FT DISULFID 675 686 BY SIMILARITY.
FT DISULFID 681 695 BY SIMILARITY.
FT DISULFID 698 709 BY SIMILARITY.
FT DISULFID 715 726 BY SIMILARITY.
FT DISULFID 721 735 BY SIMILARITY.
FT DISULFID 737 750 BY SIMILARITY.
FT DISULFID 756 767 BY SIMILARITY.
FT DISULFID 762 776 BY SIMILARITY.
FT DISULFID 778 791 BY SIMILARITY.
FT DISULFID 797 808 BY SIMILARITY.
FT DISULFID 803 817 BY SIMILARITY.
FT DISULFID 819 832 BY SIMILARITY.
FT DISULFID 838 850 BY SIMILARITY.
FT DISULFID 845 859 BY SIMILARITY.
FT DISULFID 861 873 BY SIMILARITY.
FT DISULFID 879 891 BY SIMILARITY.
FT DISULFID 885 900 BY SIMILARITY.

```

```

FT DISULFID 902 915 BY SIMILARITY.
FT DISULFID 921 933 BY SIMILARITY.
FT DISULFID 927 942 BY SIMILARITY.
FT DISULFID 944 957 BY SIMILARITY.
FT DISULFID 963 975 BY SIMILARITY.
FT DISULFID 970 984 BY SIMILARITY.
FT DISULFID 986 1000 BY SIMILARITY.
FT DISULFID 1101 1114 BY SIMILARITY.
FT DISULFID 1109 1123 BY SIMILARITY.
FT DISULFID 1125 1138 BY SIMILARITY.
FT DISULFID 1144 1155 BY SIMILARITY.
FT DISULFID 1150 1164 BY SIMILARITY.
FT DISULFID 1166 1179 BY SIMILARITY.
FT DISULFID 1298 1309 BY SIMILARITY.
FT DISULFID 1304 1318 BY SIMILARITY.
FT DISULFID 1320 1333 BY SIMILARITY.
FT DISULFID 1339 1354 BY SIMILARITY.
FT DISULFID 1349 1363 BY SIMILARITY.
FT DISULFID 1365 1378 BY SIMILARITY.
FT MOD_RES 647 HYDROXYLATION.
FT MOD_RES 810 HYDROXYLATION.
FT CARBOHYD 21 21 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 52 52 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 98 98 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 294 294 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 870 870 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 923 923 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1039 1039 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT SQ SEQUENCE 1394 AA; 152791 MW; DFFCA81A40B2C7D1 CRC64;

Query Match 63.5%; Score 33; DB 1; Length 1394;
Best Local Similarity 45.5%; Pred. No. 93;
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSYS 11
Db 399 KEICPGMGYT 409
::: | | | |

RESULT 15
RPOC_VIBCH STANDARD; PRT; 1401 AA.
AC Q9KV29;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE DNA-directed RNA polymerase beta' chain (EC 2.7.7.6) (Transcriptase
DE beta' chain) (RNA polymerase beta' subunit).
GN RPOC OR VC0329.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=El Tor N16961 / Serotype O1;
RX MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae."
RL Nature 406:477-483(2000).
CC -!- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION
CC OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS
CC SUBSTRATES (By similarity).
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate -> N diphosphate +
CC [RNA](N).
CC -!- SUBUNIT: THE ENZYME CONSISTS OF THE SIGMA CHAIN AND THE CORE

```



FT	DISULFID	998	1009	BY SIMILARITY.	CC
FT	DISULFID	1004	1018	BY SIMILARITY.	CC
FT	DISULFID	1020	1033	BY SIMILARITY.	CC
FT	DISULFID	1039	1051	BY SIMILARITY.	CC
FT	DISULFID	1046	1060	BY SIMILARITY.	CC
FT	DISULFID	1062	1074	BY SIMILARITY.	CC
FT	DISULFID	1080	1092	BY SIMILARITY.	CC
FT	DISULFID	1086	1101	BY SIMILARITY.	CC
FT	DISULFID	1103	1116	BY SIMILARITY.	CC
FT	DISULFID	1122	1134	BY SIMILARITY.	CC
FT	DISULFID	1128	1143	BY SIMILARITY.	CC
FT	DISULFID	1145	1158	BY SIMILARITY.	CC
FT	DISULFID	1164	1176	BY SIMILARITY.	CC
FT	DISULFID	1171	1185	BY SIMILARITY.	CC
FT	DISULFID	1187	1201	BY SIMILARITY.	CC
FT	DISULFID	1302	1315	BY SIMILARITY.	CC
FT	DISULFID	1310	1324	BY SIMILARITY.	CC
FT	DISULFID	1326	1339	BY SIMILARITY.	CC
FT	DISULFID	1345	1356	BY SIMILARITY.	CC
FT	DISULFID	1351	1365	BY SIMILARITY.	CC
FT	DISULFID	1367	1380	BY SIMILARITY.	CC
FT	DISULFID	1499	1510	BY SIMILARITY.	CC
FT	DISULFID	1503	1519	BY SIMILARITY.	CC
FT	DISULFID	1521	1534	BY SIMILARITY.	CC
FT	DISULFID	1540	1555	BY SIMILARITY.	CC
FT	DISULFID	1550	1564	BY SIMILARITY.	CC
FT	DISULFID	1566	1579	BY SIMILARITY.	CC
FT	MOD_RES	848	848	HYDROXYLATION (BY SIMILARITY).	CC
FT	MOD_RES	1011	1011	HYDROXYLATION (BY SIMILARITY).	CC
FT	CARBOHYD	495	495	N-LINKED (GLCNAC. . .) (POTENTIAL).	CC
FT	CARBOHYD	1071	1071	N-LINKED (GLCNAC. . .) (POTENTIAL).	CC
FT	CARBOHYD	1124	1124	N-LINKED (GLCNAC. . .) (POTENTIAL).	CC
FT	CARBOHYD	1240	1240	N-LINKED (GLCNAC. . .) (POTENTIAL).	CC
SQ	SEQUENCE	1595 AA;	173229 MW;	6A091EBA8556D8E5 CRC64;	CC
Query Match 63.5%; Score 33; DB 1; Length 1595;					
Best Local Similarity 45.5%; Pred. No. 1.le+02;					
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;					
QY	1 EEVVPXGMSYS 11				
DB	600 KEICPGMGMYT 610				
RESULT 17					
LTBL1_RAT	STANDARD; PRT; 1712 AA.				
ID	LTBL1_RAT				
AC	Q00918;				
DT	01-JUL-1993 (Rel. 26, Created)				
DT	01-JUL-1993 (Rel. 26, Last sequence update)				
DT	15-JUN-2002 (Rel. 41, Last annotation update)				
DE	Latent transforming growth factor beta binding protein 1 precursor				
DE	1) (transforming growth factor beta-1 binding protein 1) (TGF-beta1-BP-				
DE	1) (transforming growth factor beta-1 masking protein, large subunit).				
GN	LTBP1.				
OS	Rattus norvegicus (Rat).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.				
OX	NCBI_TaxID=10116;				
RP	[1]				
RP	SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.				
RX	MEDLINE=91062373; PubMed=2247454;				
RA	Tsuji T., Okada F., Yamaguchi K., Nakamura T.;				
RT	"Molecular cloning of the large subunit of transforming growth factor				
RT	type beta masking protein and expression of the mRNA in various rat				
RL	tissues.";				
RL	Proc. Natl. Acad. Sci. U.S.A. 87:8835-8839(1990).				
CC	-1- SUBUNIT: THE LARGE LATENT COMPLEX OF TGF-BETAL FROM PLATELETS IS				
CC	COMPOSED OF THE TGF-BETAL MOLECULE NONCOVALENTLY ASSOCIATED WITH				
CC	A MASKING PROTEIN CONSISTING OF A DISULFIDE-BONDED COMPLEX OF A				
CC	DIMER OF THE N-TERMINAL PROPEPTIDE OF THE TGF-BETAL PRECURSOR AND				

```

FT DISULFID 994 1005 BY SIMILARITY.
FT DISULFID 1000 1014 BY SIMILARITY.
FT DISULFID 1017 1028 BY SIMILARITY.
FT DISULFID 1034 1045 BY SIMILARITY.
FT DISULFID 1040 1054 BY SIMILARITY.
FT DISULFID 1056 1069 BY SIMILARITY.
FT DISULFID 1075 1086 BY SIMILARITY.
FT DISULFID 1081 1095 BY SIMILARITY.
FT DISULFID 1097 1110 BY SIMILARITY.
FT DISULFID 1116 1127 BY SIMILARITY.
FT DISULFID 1122 1136 BY SIMILARITY.
FT DISULFID 1138 1151 BY SIMILARITY.
FT DISULFID 1157 1169 BY SIMILARITY.
FT DISULFID 1164 1178 BY SIMILARITY.
FT DISULFID 1180 1192 BY SIMILARITY.
FT DISULFID 1198 1210 BY SIMILARITY.
FT DISULFID 1204 1219 BY SIMILARITY.
FT DISULFID 1221 1234 BY SIMILARITY.
FT DISULFID 1240 1252 BY SIMILARITY.
FT DISULFID 1246 1261 BY SIMILARITY.
FT DISULFID 1263 1276 BY SIMILARITY.
FT DISULFID 1282 1294 BY SIMILARITY.
FT DISULFID 1289 1303 BY SIMILARITY.
FT DISULFID 1305 1319 BY SIMILARITY.
FT DISULFID 1419 1432 BY SIMILARITY.
FT DISULFID 1427 1441 BY SIMILARITY.
FT DISULFID 1443 1456 BY SIMILARITY.
FT DISULFID 1462 1473 BY SIMILARITY.
FT DISULFID 1468 1482 BY SIMILARITY.
FT DISULFID 1484 1497 BY SIMILARITY.
FT DISULFID 1616 1627 BY SIMILARITY.
FT DISULFID 1622 1636 BY SIMILARITY.
FT DISULFID 1638 1651 BY SIMILARITY.
FT DISULFID 1657 1672 BY SIMILARITY.
FT DISULFID 1667 1681 BY SIMILARITY.
FT DISULFID 1683 1696 BY SIMILARITY.
FT CARBOHYD 339 339 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 370 370 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 416 416 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 612 612 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1042 1042 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1242 1242 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1357 1357 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 1712 AA; 186598 MW; 650BCEAA691FD134 CRC64;

Query Match 63.5%; Score 33; DB 1; Length 1712;
Best Local Similarity 45.5%; Pred. No. 1.1e+02;
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSYS 11
DB 718 KEICPGMGVY 728
::: |||:

RESULT 18
CGD2_RAT CGD2_RAT STANDARD; PRT; 288 AA.
AC Q04827;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE GL/S-specific cyclin D2 (Vin-1 proto-oncogene).
GN CCND2 OR VIN-1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93275661; PubMed=8502486;
RA Hanna Z., Jankowski M., Tremblay P., Jiang X.M., Milatovich A.,
RA Francke U., Jolicœur P.
RT "The Vin-1 gene, identified by provirus insertional mutagenesis, is

```

```

RT the cyclin D2.";
RL Oncogene 8:1661-1666(1993).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=95011623; PubMed=7926809;
RA Hosokawa Y., Onga T., Nakashima K.;
RT "Induction of D2 and D3 cyclin-encoding genes during promotion of the
RT GI/S transition by prolactin in rat ND2 cells.";
RL Gene 147:249-252(1994).
CC -!- FUNCTION: ESSENTIAL FOR THE CONTROL OF THE CELL CYCLE AT THE G1/S
CC (START) TRANSITION.
CC -!- SUBUNIT: INTERACTS WITH THE CDK4 AND CDK6 PROTEIN KINASES TO FORM
CC A SERINE/THREONINE KINASE HOLOENZYME COMPLEX. THE CYCLIN SUBUNIT
CC IMPARTS SUBSTRATE SPECIFICITY TO THE COMPLEX.
CC -!- SIMILARITY: BELONGS TO THE CYCLIN FAMILY. CYCLIN D SUBFAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL; L09752; AAA41010.1; .
EMBL; D16308; BAA03815.1; .
InterPro; IPR004366; Cyclin.
InterPro; IPR004367; Cyclin_Cterm.
DR Pfam; PF001134; cyclin; 1.
DR Pfam; PF02984; cyclin_C; 1.
DR SMART; SM00385; cyclin; 1.
DR PROSITE; PS00292; CYCLINS; 1.
DR Cyclin; Cell cycle; Cell division; Multigene family; Proto-oncogene.
KW CONFLICT 68 68 E -> G (IN REF. 2).
FT CONFLICT 104 104 C -> V (IN REF. 2).
FT CONFLICT 232 232 T -> A (IN REF. 2).
SQ SEQUENCE 288 AA; 32826 MW; 4B522BF4E9835FC1 CRC64;

Query Match 61.5%; Score 32; DB 1; Length 288;
Best Local Similarity 60.0%; Pred. No. 29;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10
DB 73 EEVFPAMNY 82
||| | | |

RESULT 19
CGD2_HUMAN CGD2_HUMAN STANDARD; PRT; 289 AA.
AC P30279; Q13955;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE GL/S-specific cyclin D2.
GN CCND2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92347851; PubMed=1386336;
RA Xiong Y., Menninger J., Beach D., Ward D.C.;
RT "Molecular cloning and chromosomal mapping of CCND genes encoding
RT human D-type cyclins.";
RL Genomics 13:575-584(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=93205384; PubMed=8455931;
RA Palmero I., Holder A., Sinclair A.J., Dickson C., Peters G.;
RT "Cyclins D1 and D2 are differentially expressed in human B-lymphoid
RT cell lines.";

```



OncoGene 8:1049-1054(1993).

[3]  
RN SEQUENCE FROM N.A.  
RA Miyajima N.;  
RL Submitted (MAR-1993) to the EMBL/GenBank/DBJ databases.  
[4]  
RN SEQUENCE FROM N.A.  
RC TISSUE-Bone marrow;  
RA Strausberg R.;  
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
[5]  
RN SEQUENCE OF 1-240 FROM N.A.  
RC TISSUE-Placenta;  
RA Inaba T., Matsushima H., Valentine M., Rousset M.F., Sherr C.J.,  
RA Look A.T.;  
RT "Genomic organization, chromosomal localization, and independent  
RT expression of human cyclin D genes.";  
RL Genomics 13:565-574(1992).  
CC -!- FUNCTION: ESSENTIAL FOR THE CONTROL OF THE CELL CYCLE AT THE G1/S  
CC (START) TRANSITION.  
CC -!- SUBUNIT: INTERACTS WITH THE CDK4 AND CDK6 PROTEIN KINASES TO FORM  
CC A SERINE/THREONINE KINASE HOLOENZYME COMPLEX. THE CYCLIN SUBUNIT  
CC IMPARTS SUBSTRATE SPECIFICITY TO THE COMPLEX.  
CC -!- SIMILARITY: BELONGS TO THE CYCLIN FAMILY. CYCLIN D SUBFAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL; M50813; AAA51926.1; -;  
DR EMBL; X68452; CAA48493.1; -;  
DR EMBL; D13639; BAA02802.1; -;  
DR EMBL; BC010958; AAH10958.1; -;  
DR EMBL; M88083; AAA51928.1; -;  
DR EMBL; M88080; AAA51928.1; JOINED.  
DR EMBL; M88081; AAA51928.1; JOINED.  
DR EMBL; M88082; AAA51928.1; JOINED.  
DR PIR; A42822; A42822.  
DR PIR; S26580; S26580.  
DR Genew; HGNC:1583; CCND2.  
DR MIM; 123833; -;  
DR InterPro; IPR004366; Cyclin.  
DR InterPro; IPR004367; Cyclin\_Cterm.  
DR Pfam; PF00134; cyclin; 1.  
DR Pfam; PF02984; cyclin\_C; 1.  
DR SMART; SM00385; CYCLIN; 1.  
DR PROSITE; PS00292; CYCLINS; 1.  
KW Cyclin; Cell cycle; Cell division; Multigene family.  
FT CONFLICT 166 167 KL -> NV (IN REF. 5).  
FT CONFLICT 224 224 T -> H (IN REF. 5).  
SQ SEQUENCE 289 AA; 33067 MW; E4E5FEF476D76D90 CRC64;  
  
Query Match 61.5%; Score 32; DB 1; Length 289;  
Best Local Similarity 60.0%; Pred. No. 29;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
  
QY 1 EEVVPXGMSY 10  
Db ||| | | | | |  
74 EEVFPPLANNY 83  
  
RESULT 20  
CGD2\_MOUSE  
ID CGD2\_MOUSE STANDARD; PRT; 289 AA.  
AC P30280;  
DT 01-APR-1993 (Rel. 25, Created)  
DT 01-APR-1993 (Rel. 25, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)

G1/S-specific cyclin D2.  
CCND2 OR CYL-2.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
[1]  
RN SEQUENCE FROM N.A.  
RX MEDLINE=92196134; PubMed=1372445;  
RA Kiyokawa H., Busquets X., Powell C.T., Ngo L., Rifkin R.A.,  
RA Marks P.A.;  
RT "Cloning of a D-type cyclin from murine erythroleukemia cells.";  
RL Proc. Natl. Acad. Sci. U.S.A. 89:2444-2447(1992).  
[2]  
RN SEQUENCE FROM N.A.  
RX MEDLINE=91235305; PubMed=1827757;  
RA Matsushima H., Rousset M.F., Ashmun R.A., Sherr C.J.;  
RT "Colony-stimulating factor 1 regulates novel cyclins during the G1  
RT phase of the cell cycle.";  
RL Cell 65:701-713(1991).  
CC -!- FUNCTION: ESSENTIAL FOR THE CONTROL OF THE CELL CYCLE AT THE G1/S  
CC (START) TRANSITION.  
CC -!- SUBUNIT: INTERACTS WITH THE CDK4 AND CDK6 PROTEIN KINASES TO FORM  
CC A SERINE/THREONINE KINASE HOLOENZYME COMPLEX. THE CYCLIN SUBUNIT  
CC IMPARTS SUBSTRATE SPECIFICITY TO THE COMPLEX.  
CC -!- SIMILARITY: BELONGS TO THE CYCLIN FAMILY. CYCLIN D SUBFAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL; M83749; AAA37519.1; -;  
DR EMBL; M86182; AAA37503.1; -;  
DR PIR; B40035; B40035.  
DR PIR; A41984; A41984.  
DR MGD; MGI:88314; Ccnd2.  
DR InterPro; IPR004366; Cyclin.  
DR InterPro; IPR004367; Cyclin\_Cterm.  
DR Pfam; PF00134; cyclin; 1.  
DR Pfam; PF02984; cyclin\_C; 1.  
DR SMART; SM00385; CYCLIN; 1.  
DR PROSITE; PS00292; CYCLINS; 1.  
KW Cyclin; Cell cycle; Cell division; Multigene family.  
SQ SEQUENCE 289 AA; 32897 MW; 58F322771DD1DA3D CRC64;  
  
Query Match 61.5%; Score 32; DB 1; Length 289;  
Best Local Similarity 60.0%; Pred. No. 29;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
  
QY 1 EEVVPXGMSY 10  
Db ||| | | | | |  
73 EEVFPPLANNY 82  
  
RESULT 21  
CGD1\_BRARE  
ID CGD1\_BRARE STANDARD; PRT; 291 AA.  
AC Q90459;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE G1/S-specific cyclin D1.  
GN CYCD1.  
OS Brachydanio rerio (Zebrafish) (Danio rerio).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
OC Cyprinidae; Danio.  
OX NCBI\_TaxID=7955;  
RN [1]

RP SEQUENCE FROM N.A.  
RA MEDLINE-96138542; PubMed-8547308;  
RX Yarden A., Salomon D., Geiger B.;  
RT "zebrafish cyclin D1 is differentially expressed during early  
RL embryogenesis.";  
RL Biochim. Biophys. Acta 1264:257-260(1995).  
CC -1- FUNCTION: ESSENTIAL FOR THE CONTROL OF THE CELL CYCLE AT THE G1/S  
CC (START) TRANSITION.  
CC -1- SUBUNIT: INTERACTS WITH THE CDK4 AND CDK6 PROTEIN KINASES TO FORM  
CC A SERINE/THREONINE KINASE HOLOENZYME COMPLEX. THE CYCLIN SUBUNIT  
CC IMPARTS SUBSTRATE SPECIFICITY TO THE COMPLEX.  
CC -1- SIMILARITY: BELONGS TO THE CYCLIN FAMILY. CYCLIN D SUBFAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
DR EMBL; X87581; CAA60885.1; -  
DR ZFIN; ZDB-GENE-980526-176; cycd1.  
DR InterPro; IPR004366; Cyclin\_Cterm.  
DR InterPro; IPR004367; Cyclin\_Cterm.  
DR Pfam; PF00134; cyclin; 1.  
DR Pfam; PF02984; cyclin\_C; 1.  
DR SMART; SM00385; CYCLIN; 1.  
DR PROSITE; PS00292; CYCLINS; 1.  
KW Cyclin; Cell cycle; Cell division.  
SQ SEQUENCE 291 AA; 33067 MW; FA5274CB1B4D5EF CRC64;  
  
Query Match 61.5%; Score 32; DB 1; Length 291;  
Best Local Similarity 60.0%; Pred. No. 29;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
  
QY 1 EEVVPXGMSY 10  
||| | |  
DB 75 EEVFPPLAMNY 84  
  
RESULT 22  
CGDL\_XENLA STANDARD; PRT; 291 AA.  
AC P30755;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE G1/S-specific cyclin D1.  
GN CCND1  
OS Xenopus laevis (African clawed frog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;  
OC Xenopodinae; Xenopus.  
OX NCBI\_TaxID=8355;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Cockerill M.J., Hunt T.;  
RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: ESSENTIAL FOR THE CONTROL OF THE CELL CYCLE AT THE G1/S  
CC (START) TRANSITION.  
CC -1- SUBUNIT: INTERACTS WITH THE CDK4 AND CDK6 PROTEIN KINASES TO FORM  
CC A SERINE/THREONINE KINASE HOLOENZYME COMPLEX. THE CYCLIN SUBUNIT  
CC IMPARTS SUBSTRATE SPECIFICITY TO THE COMPLEX.  
CC -1- SIMILARITY: BELONGS TO THE CYCLIN FAMILY. CYCLIN D SUBFAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC

CC EMBL; X89475; CAA61664.1; -  
DR InterPro; IPR004366; Cyclin.  
DR InterPro; IPR004367; Cyclin\_Cterm.  
DR Pfam; PF00134; cyclin; 1.  
DR Pfam; PF02984; cyclin\_C; 1.  
DR SMART; SM00385; CYCLIN; 1.  
DR PROSITE; PS00292; CYCLINS; 1.  
KW Cyclin; Cell cycle; Cell division; Multigene family.  
SQ SEQUENCE 291 AA; 32953 MW; A4747C5BD1679087 CRC64;  
  
Query Match 61.5%; Score 32; DB 1; Length 291;  
Best Local Similarity 60.0%; Pred. No. 29;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
  
QY 1 EEVVPXGMSY 10  
||| | |  
DB 73 EEVFPPLAMNY 82  
  
RESULT 23  
CGD2\_CHICK STANDARD; PRT; 291 AA.  
AC P49706;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE G1/S-specific cyclin D2.  
GN CCND2.  
OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
OC Gallus.  
OX NCBI\_TaxID=9031;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-96144302; PubMed-8566807;  
RA Li H., Grenet J., Kidd V.J.;  
RT "Structure and gene expression of avian cyclin D2.";  
RL Gene 167:341-342(1995).  
CC -1- FUNCTION: ESSENTIAL FOR THE CONTROL OF THE CELL CYCLE AT THE G1/S  
CC (START) TRANSITION.  
CC -1- SUBUNIT: INTERACTS WITH THE CDK4 AND CDK6 PROTEIN KINASES TO FORM  
CC A SERINE/THREONINE KINASE HOLOENZYME COMPLEX. THE CYCLIN SUBUNIT  
CC IMPARTS SUBSTRATE SPECIFICITY TO THE COMPLEX.  
CC -1- SIMILARITY: BELONGS TO THE CYCLIN FAMILY. CYCLIN D SUBFAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
DR EMBL; U28980; AAA96955.1; -  
DR InterPro; IPR004366; Cyclin.  
DR InterPro; IPR004367; Cyclin\_Cterm.  
DR Pfam; PF00134; cyclin; 1.  
DR Pfam; PF02984; cyclin\_C; 1.  
DR SMART; SM00385; CYCLIN; 1.  
DR PROSITE; PS00292; CYCLINS; 1.  
KW Cyclin; Cell cycle; Cell division; Multigene family.  
SQ SEQUENCE 291 AA; 33163 MW; 43A7E646AEAF3109 CRC64;  
  
Query Match 61.5%; Score 32; DB 1; Length 291;  
Best Local Similarity 60.0%; Pred. No. 29;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
  
QY 1 EEVVPXGMSY 10  
||| | |  
DB 74 EEVFPPLAMNY 83

## RESULT 24

CGD2\_XENLA  
ID CGD2\_XENLA STANDARD; PRT; 291 AA.  
AC P53782;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE G1/S-specific cyclin D2.  
GN CCND2.  
OS Xenopus laevis (African clawed frog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;  
OC Xenopodinae; Xenopus.  
OX NCBI\_TaxID=8355;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Cockerill M.J., Hunt T.;  
RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97380591; PubMed=9237366;  
RA Taieb F., Jessus C.;  
RT "Xenopus cyclin D2: cloning and expression in oocytes and during  
early development.";  
RL Biol. Cell 88:99-111(1996).  
CC -!- FUNCTION: ESSENTIAL FOR THE CELL CYCLE AT THE G1/S  
(START) TRANSITION.  
CC -!- SUBUNIT: INTERACTS WITH THE CDK4 AND CDK6 PROTEIN KINASES TO FORM  
A SERINE/THREONINE KINASE HOLOENZYME COMPLEX. THE CYCLIN SUBUNIT  
IMPARTS SUBSTRATE SPECIFICITY TO THE COMPLEX.  
CC -!- SIMILARITY: BELONGS TO THE CYCLIN FAMILY. CYCLIN D SUBFAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL; X89476; CAA61665.1; -;  
DR EMBL; X83503; CAA58493.1; -;  
DR InterPro; IPR004366; Cyclin.  
DR InterPro; IPR004367; Cyclin\_Cterm.  
DR Pfam; PF00134; cyclin; 1.  
DR Pfam; PF02984; cyclin\_C; 1.  
DR SMART; SM00385; CYCLIN; 1.  
DR PROSITE; PS00292; CYCLINS; 1.  
KW Cyclin; Cell cycle; Cell division; Multigene family.  
SQ SEQUENCE 291 AA; 32959 MW; 9A290F04F1531E89 CRC64;  
Query Match 61.5%; Score 32; DB 1; Length 291;  
Best Local Similarity 60.0%; Pred. No. 29;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
QY 1 EEVVPXGMSY 10  
Db 74 EEVFPAMNYY 83

## RESULT 25

CGD1\_CHICK  
ID CGD1\_CHICK STANDARD; PRT; 292 AA.  
AC P53169;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE G1/S-specific cyclin D1.  
GN CCND1.  
OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
OC

Gallus.  
OX NCBI\_TaxID=9031;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Li H., Lahti J.M., Kidd V.J.;  
RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: ESSENTIAL FOR THE CONTROL OF THE CELL CYCLE AT THE G1/S  
(START) TRANSITION. INTERACTS WITH THE CDK4 AND CDK6 PROTEIN  
KINASES (BY SIMILARITY).  
CC -!- SIMILARITY: BELONGS TO THE CYCLIN FAMILY. CYCLIN D SUBFAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL; U40844; AAA83271.1; -;  
DR InterPro; IPR004366; Cyclin.  
DR InterPro; IPR004367; Cyclin\_Cterm.  
DR Pfam; PF00134; cyclin; 1.  
DR Pfam; PF02984; cyclin\_C; 1.  
DR SMART; SM00385; CYCLIN; 1.  
DR PROSITE; PS00292; CYCLINS; 1.  
KW Cyclin; Cell cycle; Cell division; Multigene family.  
SQ SEQUENCE 292 AA; 33263 MW; 7B543029DB45A67D CRC64;  
Query Match 61.5%; Score 32; DB 1; Length 292;  
Best Local Similarity 60.0%; Pred. No. 30;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
QY 1 EEVVPXGMSY 10  
Db 75 EEVFPAMNYY 84  
Search completed: June 10, 2003, 13:40:16  
Job time : 5.5 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2003, 13:30:25 ; Search time 25.7857 Seconds

(without alignments)  
87.898 Million cell updates/sec

Title: US-09-909-164-5  
Perfect score: 52  
Sequence: 1 EHVVPXGMSYS 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 75 summaries

#### Database :

SPREMBL\_21.\*  
1: sp.archaea.\*  
2: sp.bacteria.\*  
3: sp.fungi.\*  
4: sp.human.\*  
5: sp.invertebrate.\*  
6: sp.mammal.\*  
7: sp.mhc.\*  
8: sp.organella.\*  
9: sp.phage.\*  
10: sp.plant.\*  
11: sp.rodent.\*  
12: sp.virus.\*  
13: sp.vertebrate.\*  
14: sp.unclassified.\*  
15: sp.virus.\*  
16: sp.bacteriopl.\*  
17: sp.archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	38	73.1	387	16 Q98FX1	Q98fx1 rhizobium 1
2	38	73.1	1063	16 Q8RG86	Q8rg86 fusobacteri
3	38	73.1	3472	1 Q74056	Q74056 cenarchaeum
4	37	71.2	840	3 Q9URY8	Q9ury8 schizosacch
5	36	69.2	471	11 Q8R126	Q8rl26 mus musculu
6	36	69.2	484	11 Q8VD18	Q8vd18 mus musculu
7	35	67.3	225	10 Q40129	Q40129 lycopersico
8	35	67.3	425	5 Q9XVK4	Q9xvk4 caenorhabdi
9	35	67.3	556	4 Q43733	Q43733 homo sapien
10	35	67.3	583	5 Q9BHA5	Q9bha5 plasmodium
11	35	67.3	583	5 Q9BH83	Q9bh83 plasmodium
12	35	67.3	670	11 Q01487	Q01487 rattus norv
13	35	67.3	749	16 Q9PDM6	Q9pdm6 xylella fas
14	35	67.3	1902	4 Q14122	Q14122 homo sapien
15	34	65.4	156	3 Q12479	Q12479 saccharomyc
16	34	65.4	219	17 Q971S2	Q971s2 sulfolobus

#### RESULT 1

Q98FX1 PRELIMINARY; PRG; 387 AA.  
ID Q98FX1  
AC Q98FX1  
DT 01-OCT-2001 (TrEMBLrel. 18, Created)  
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)  
DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)  
DE Hippurate hydrolase.  
GN MLR3583.

#### ALIGNMENTS

34	65.4	252	17	O28342	O28342 archaeoglob
34	65.4	290	4	Q96MU1	Q96mu1 homo sapien
34	65.4	387	16	Q92MD6	Q92md6 rhizobium m
34	65.4	541	16	Q98BP5	Q98bp5 rhizobium l
34	65.4	544	16	Q9PQD2	Q9pqd2 ureaplasma
34	65.4	842	3	Q9URR4	Q9urr4 penicillium
34	65.4	1049	16	Q8XT05	Q8xt05 ralstonia s
34	65.4	1499	4	Q96914	Q96914 homo sapien
33	63.5	143	17	Q8TX62	Q8tx62 methanopyru
33	63.5	162	11	Q9CXQ4	Q9cxq4 mus musculu
33	63.5	165	17	Q8B330	Q8b330 archaeoglob
33	63.5	193	2	Q8VUA8	Q8vua8 lactococcus
33	63.5	209	16	Q8RE56	Q8re56 fusobacteri
33	63.5	284	16	P74187	P74187 synchocyst
33	63.5	288	10	Q9M3C0	Q9m3c0 arabidopsis
33	63.5	326	12	Q9Q9Q9	Q9q9q9 soil-borne
33	63.5	326	12	Q9Q9Q5	Q9qq95 soil-borne
33	63.5	326	12	Q9QCE7	Q9qce7 soil-borne
33	63.5	326	12	Q9DJG4	Q9djg4 soil-borne
33	63.5	326	12	Q91DN1	Q91dn1 soil-borne
33	63.5	326	12	Q9Q9Q7	Q9q9q7 soil-borne
33	63.5	327	12	Q06360	Q06360 soil-borne
33	63.5	332	10	Q9FN14	Q9fn14 arabidopsis
33	63.5	368	16	Q9X0U3	Q9x0u3 thermotoga
33	63.5	393	5	Q9V914	Q9v914 drosophila
33	63.5	41	4	Q96CS0	Q96cs0 homo sapien
33	63.5	479	4	Q96CS0	Q96cs0 mus musculu
33	63.5	548	11	Q9D2X9	Q9d2x9 mus musculu
33	63.5	584	16	Q8R8K6	Q8r8k6 thermoanaer
33	63.5	648	4	Q96MB2	Q96mb2 homo sapien
33	63.5	653	16	Q9KVE3	Q9kve3 vibrio chol
33	63.5	676	5	Q9VA55	Q9va55 drosophila
33	63.5	676	5	Q8T827	Q8t827 drosophila
33	63.5	678	12	Q9ELX6	Q9elx6 cercopithec
33	63.5	746	3	Q9URR3	Q9urr3 penicillium
33	63.5	791	4	Q9H2K5	Q9h2k5 homo sapien
33	63.5	793	4	Q9H2K6	Q9h2k6 homo sapien
33	63.5	844	11	Q922D4	Q922d4 mus musculu
33	63.5	1028	16	Q8YJ11	Q8yj11 bruceella me
33	63.5	1152	16	Q9CC95	Q9cc95 mycobacteri
33	63.5	1305	5	Q9V7C7	Q9v7c7 drosophila
33	63.5	1394	4	Q8TD95	Q8td95 homo sapien
33	63.5	1442	17	Q96YH5	Q96yh5 sulfolobus
33	63.5	1548	10	Q65531	Q65531 arabidopsis
33	63.5	1713	11	Q88349	Q88349 mus musculu
33	61.5	84	16	Q97DE7	Q97de7 clostridium
32	61.5	103	11	Q9DOH9	Q9doh9 mus musculu
32	61.5	108	1	Q9UX33	Q9ux33 sulfolobus
32	61.5	153	13	P79919	P79919 xenopus lae
32	61.5	156	11	Q9D8L9	Q9d8l9 mus musculu
32	61.5	174	10	Q9M3T4	Q9m3t4 betula verr
32	61.5	190	13	Q57481	Q57481 stizostedio
32	61.5	191	11	Q99NB4	Q99nb4 rattus norv
32	61.5	200	17	Q97CD0	Q97cd0 thermoplasma
32	61.5	207	2	Q47284	Q47284 escherichia
32	61.5	234	2	O32330	O32330 clostridium
32	61.5	236	10	Q9SXF1	Q9sxf1 arabidopsis
32	61.5	240	11	Q9DB09	Q9db09 mus musculu
32	61.5	243	12	Q91EW1	Q91ew1 cydia pomon
32	61.5	279	17	Q9Y824	Q9y824 aeropyrum p

```

OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpō S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti.";
RL DNA Res. 7:331-338(2000).
DR EMBL; AP003002; BAB50445.1; -;
DR InterPro; IPR002933; Peptidase_M20.
DR Pfam; PF01546; Peptidase_M20; 1.
DR Hydrolase; Complete proteome.
KW SEQUENCE 387 AA; 41180 MW; 131BFF8E64306829 CRC64;
SQ
Query Match 73.1%; Score 38; DB 16; Length 387;
Best Local Similarity 60.0%; Pred. No. 9.8;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 1 EEVVPXGMSY 10
Db :|:|:|:|
367 DEAIPIHGMSY 376

RESULT 2
Q8RG86 PRELIMINARY; PRT; 1063 AA.
ID Q8RG86
AC Q8RG86;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Carbamoyl-phosphate synthase large chain (EC 6.3.5.5).
GN FN0422.
OS Fusobacterium nucleatum (subsp. nucleatum).
OC Bacteria; Fusobacteria; Fusobacterium.
OX NCBI_TaxID=76856;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 25586;
RX MEDLINE=21886394; PubMed=11899109;
RA Kapatal V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A.,
RA Bhattacharyya A., Bartman A., Gardner W., Grechkin G., Zhu L.,
RA Vasieva O., Chu L., Kogan Y., Chaga O., Goltsman E., Bernal A.,
RA Larsen N., D'Souza M., Walunas T., Pusch G., Haselkorn R.,
RA Fonstein M., Kyriades N., Overbeek R.;
RT "Genome sequence and analysis of the oral bacterium Fusobacterium
RT nucleatum strain ATCC 25586.";
RL J. Bacteriol. 184:2005-2018(2002).
DR EMBL; AE010554; AAL94625.1; -;
KW Ligase; Complete proteome.
SQ SEQUENCE 1063 AA; 118008 MW; 39700E10B7CCE411 CRC64;
Query Match 73.1%; Score 38; DB 16; Length 1063;
Best Local Similarity 60.0%; Pred. No. 29;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 2 EVVPXGMSYS 11
Db :|:|:|:|
195 EIVPGLNYS 204

RESULT 3
O74056 PRELIMINARY; PRT; 3472 AA.
ID O74056
AC O74056;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)

```

```

DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical 367.1 kDa protein.
OS Cenarchaeum symbiosum.
OC Archaea; Crenarchaeota; Cenarchaeum.
OX NCBI_TaxID=46770;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=B;
RX MEDLINE=98422450; PubMed=9748430;
RA Schleper C., DeLong E.F., Preston C.M., Feldman R.A., Wu K.Y.,
RA Swanson R.V.;
RT "Genomic analysis reveals chromosomal variation in natural populations
RT of the uncultured psychrophilic archaeon Cenarchaeum symbiosum.";
RL J. Bacteriol. 180:5003-5009(1998).
CC -1- SIMILARITY: CONTAINS 5 WD REPEATS (TRP-ASP DOMAINS).
DR EMBL; AF083072; AAC62699.1; -;
DR InterPro; IPR000515; BPD_transp.
DR InterPro; IPR001680; WD40.
DR Pfam; PF00400; WD40; 4.
DR SMART; SM00320; WD40; 2.
DR PROSITE; PS00402; BPD_TRANSF_INN_MEMBR; UNKNOWN_1.
KW Hypothetical protein; Repeat; WD repeat.
SQ SEQUENCE 3472 AA; 367058 MW; 37F80707030F9355 CRC64;
Query Match 73.1%; Score 38; DB 1; Length 3472;
Best Local Similarity 54.5%; Pred. No. 1e+02;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
QY 1 EEVVPXGMSYS 11
Db :|:|:|:|
2294 EDVIPRGISFS 2304

RESULT 4
Q9URY8 PRELIMINARY; PRT; 840 AA.
ID Q9URY8
AC Q9URY8;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Probable sulfate permease.
GN SPAC869.05C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972H-;
RA Hunt C., Aves S., McDougall R.C., Rajandream M.A., Barrell B.G.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL132779; CAB60015.1; -;
DR InterPro; IPR002645; STAS.
DR InterPro; IPR001902; Sulfate_transp.
DR Pfam; PF01740; STAS; 1.
DR Pfam; PF00916; Sulfate_transp; 1.
DR TIGRFAMs; TIGR00815; sulp; 1.
SQ SEQUENCE 840 AA; 93517 MW; ED4833E162B69077 CRC64;
Query Match 71.2%; Score 37; DB 3; Length 840;
Best Local Similarity 77.8%; Pred. No. 37;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 3 VVPXGMSYS 11
Db :|:|:|:|
135 VVPOGMSYA 143

RESULT 5
Q8R126 PRELIMINARY; PRT; 471 AA.
ID Q8R126

```

AC Q8R126;  
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
 DE Hypothetical 54.5 kDa protein (Fragment).  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=LIVER;  
 RA Strausberg R.;  
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC025810; AAH25810.1; -  
 KW Hypothetical protein.  
 FT NON\_TER  
 SQ SEQUENCE 471 AA; 54506 MW; E0DA685C374A9760 CRC64;  
 Query Match 69.2%; Score 36; DB 11; Length 471;  
 Best Local Similarity 60.0%; Pred. No. 32;  
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
 QY 2 EVVPXGMSYS 11  
 Db 226 EVIPAGASYN 235  
 RESULT 6  
 Q8VD18 PRELIMINARY; PRT; 484 AA.  
 AC Q8VD18;  
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
 DE Similar to glioma tumor suppressor candidate region gene 2.  
 GN AW536441.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=SALIVARY GLAND;  
 RA Strausberg R.;  
 RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC017637; AAH17637.1; -  
 DR MGD; MGI:2138595; AW536441.  
 SQ SEQUENCE 484 AA; 55835 MW; BBB45F3B4BE02A36 CRC64;  
 Query Match 69.2%; Score 36; DB 11; Length 484;  
 Best Local Similarity 60.0%; Pred. No. 33;  
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
 QY 2 EVVPXGMSYS 11  
 Db 239 EVIPAGASYN 248  
 RESULT 7  
 Q40129 PRELIMINARY; PRT; 225 AA.  
 AC Q40129;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DE Hypothetical 25.2 kDa protein precursor.  
 OS Lycopersicon esculentum (Tomato).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.  
 OX NCBI\_TaxID=4081;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RC STRAIN-VF36; TISSUE=PISTIL;  
 RX MEDLINE=95375233; PubMed=7647301;  
 RA Milligan S.B.; Gasser C.S.;  
 RT "Nature and regulation of pistil-expressed genes in tomato."  
 RL Plant Mol. Biol. 28:691-711(1995).  
 DR EMBL; U20592; AAA80497.1; -  
 DR InterPro; IPR002160; Kunitz\_legume.  
 DR Pfam; PF00197; Kunitz\_legume; 1.  
 DR ProDom; PD000891; Kunitz\_legume; 1.  
 DR SMART; SM00452; STI; 1.  
 DR PROSITE; PS00283; SOYBEAN\_KUNITZ; UNKNOWN\_1.  
 KW Hypothetical protein; Signal.  
 FT SIGNAL 1 20 POTENTIAL.  
 FT CHAIN 21 225 UNKNOWN.  
 SQ SEQUENCE 225 AA; 25188 MW; 1074C261D20CFDAD CRC64;  
 Query Match 67.3%; Score 35; DB 10; Length 225;  
 Best Local Similarity 54.5%; Pred. No. 24;  
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 EVVPXGMSYS 11  
 Db 32 DEVVPGKTYA 42  
 RESULT 8  
 Q9XVK4 PRELIMINARY; PRT; 425 AA.  
 AC Q9XVK4;  
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 DE R10D12.10 protein.  
 GN R10D12.10.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Percy C.M.;  
 RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99069613; PubMed=9851916;  
 RA none;  
 RT "Genome sequence of the nematode C.elegans: A platform for  
 RL investigating biology."  
 RL Science 282:2012-2018(1998).  
 DR EMBL; Z81109; CAB03241.1; -  
 DR InterPro; IPR000719; Euk\_pkinase.  
 DR Pfam; PF00069; pkinase; 1.  
 DR ProDom; PD000001; Euk\_pkinase; 1.  
 DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
 KW ATP-binding; Transferase.  
 SQ SEQUENCE 425 AA; 49410 MW; 5D96E29B08C8E9D6 CRC64;  
 Query Match 67.3%; Score 35; DB 5; Length 425;  
 Best Local Similarity 50.0%; Pred. No. 48;  
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 EVVPXGMSY 10  
 Db 335 EQUIVPGGLQY 344  
 RESULT 9  
 Q43733 PRELIMINARY; PRT; 556 AA.  
 ID Q43733  
 AC Q43733;  
 DT 01-JUN-1998 (TrEMBLrel. 06, Created)  
 DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)

DT 01-JUN-1998 (TReMBLrel. 06, Last annotation update)  
 DE DNA binding protein (Fragment).  
 GN DJ451B15.2.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN RP SEQUENCE FROM N.A.  
 RA Tubby B.;  
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; Z98050; CAB10847.1; -.  
 FT NON\_TER 1  
 SQ SEQUENCE 556 AA; 59059 MW; EC800E4033FB2528 CRC64;  
  
 Query Match 67.3%; Score 35; DB 4; Length 556;  
 Best Local Similarity 66.7%; Pred. No. 64;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
  
 QY 3 VVPXGMSYS 11  
 Db 244 VVPAGLYS 252  
  
 RESULT 10  
 Q9BHA5 PRELIMINARY; PRT; 583 AA.  
 ID Q9BHA5;  
 AC Q9BHA5;  
 DT 01-JUN-2001 (TReMBLrel. 17, Created)  
 DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)  
 DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)  
 DE Choline transporter.  
 GN SCTL.  
 OS Plasmodium falciparum.  
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
 OX NCBI\_TaxID=5833;  
 RN RP SEQUENCE FROM N.A.  
 RA Ben Mamoun C., Gluzman I.Y., Goldberg D.E.;  
 RL "Plasmodium falciparum choline transporter (PfSCTL) gene."  
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY007372; AAK14816.1; -.  
 DR EMBL; AY007375; AAG17947.1; -.  
 DR InterPro; IPR002123; Acyltransferase.  
 DR Pfam; PF01553; Acyltransferase; 1.  
 SQ SEQUENCE 583 AA; 66917 MW; 2B2BFAE3E395E049 CRC64;  
  
 Query Match 67.3%; Score 35; DB 5; Length 583;  
 Best Local Similarity 55.6%; Pred. No. 67;  
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
  
 QY 3 VVPXGMSYS 11  
 Db 227 IIPVGLSYS 235  
  
 RESULT 11  
 Q9BH83 PRELIMINARY; PRT; 583 AA.  
 ID Q9BH83;  
 AC Q9BH83;  
 DT 01-JUN-2001 (TReMBLrel. 17, Created)  
 DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)  
 DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)  
 DE Choline transporter.  
 GN SCTL.  
 OS Plasmodium falciparum.  
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
 OX NCBI\_TaxID=5833;  
 RN RP SEQUENCE FROM N.A.  
 RA Ben Mamoun C., Gluzman I.Y., Goldberg D.E.;  
 RL "Plasmodium falciparum choline transporter (PfSCTL) gene."  
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AY007374; AAK14818.1; -.  
 DR EMBL; AY007373; AAK14817.1; -.  
 DR InterPro; IPR002123; Acyltransferase.  
 DR Pfam; PF01553; Acyltransferase; 1.  
 SQ SEQUENCE 583 AA; 66918 MW; 4DF83D7530527474 CRC64;  
  
 Query Match 67.3%; Score 35; DB 5; Length 583;  
 Best Local Similarity 55.6%; Pred. No. 67;  
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
  
 QY 3 VVPXGMSYS 11  
 Db 227 IIPVGLSYS 235  
  
 RESULT 12  
 Q01487 PRELIMINARY; PRT; 670 AA.  
 ID Q01487;  
 AC Q01487;  
 DT 01-NOV-1996 (TReMBLrel. 01, Created)  
 DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)  
 DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)  
 DE DNA-binding protein AT-BP2 (ALPHA1-antitrypsin promoter binding protein 2) (Fragment).  
 DE Rattus norvegicus (Rat).  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN RP SEQUENCE FROM N.A.  
 RC STRAIN-SPRAGUE-DAWLEY; TISSUE=THYROID;  
 RX MEDLINE=91187610; PubMed=1901405;  
 RA Mitchelmore C., Traboni C., Cortese R.;  
 RT "Isolation of two cDNAs encoding zinc finger proteins which bind to the alpha 1-antitrypsin promoter and to the major histocompatibility complex class I enhancer."  
 RT Nucleic Acids Res. 19:141-147(1991).  
 RL -I- FUNCTION: BINDS TO THE ALPHA1-ANTITRYPSIN PROMOTER, TO THE KAPPA IMMUNOGLOBULIN GENE ENHANCER, AND TO THE MAJOR HISTOCOMPATIBILITY COMPLEX CLASS I ENHANCER; IT PLAYS A ROLE AS TRANSCRIPTIONAL REGULATOR. CONCERNING THE ALPHA1-ANTITRYPSIN EXPRESSION IT MAY ACT AS A NEGATIVE REGULATOR BY INTERFERING AT THE ALPHA1-ANTITRYPSIN PROMOTER WITH ANOTHER NUCLEAR PROTEIN, CALLED LF-B.  
 CC -I- SUBCELLULAR LOCATION: NUCLEAR.  
 CC -I- MISCELLANEOUS: ACIDIC AMINO ACID REGION MAY BE INVOLVED IN TRANSCRIPTIONAL ACTIVATION.  
 CC -I- SIMILARITY: HIGHLY SIMILAR TO AT-BP1 AND TO THE PRDII-BF1 FACTOR.  
 DR EMBL; X54250; CAA38151.1; -.  
 DR HSP; P15822; 1BBO.  
 DR InterPro; IPR000822; znf\_C2H2.  
 DR Pfam; PF00096; zf\_C2H2; 2.  
 DR SMART; SM00355; znf\_C2H2; 2.  
 DR PROSITE; PS00028; ZINC\_FINGER\_C2H2\_1; 2.  
 DR PROSITE; PS0157; ZINC\_FINGER\_C2H2\_2; 2.  
 KW Zinc-finger; Nuclear protein; DNA-binding; Transcription regulation;  
 KW Metal-binding; Multigene family.  
 FT NON\_TER 1  
 FT DOMAIN 54 104 ZINC-FINGERS.  
 FT DOMAIN 140 160 ACIDIC.  
 FT ZN\_FING 54 74 C(2)H(2) CLASS.  
 FT ZN\_FING 82 104 C(2)H(2) CLASS.  
 SQ SEQUENCE 670 AA; 71233 MW; CDD2324152590C17 CRC64;  
  
 Query Match 67.3%; Score 35; DB 11; Length 670;  
 Best Local Similarity 66.7%; Pred. No. 78;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
  
 QY 3 VVPXGMSYS 11  
 Db 376 VVPAGLYS 384  
  
 RESULT 13



Q9PDM6  
ID Q9PDM6 PRELIMINARY; PRT; 749 AA.  
AC Q9PDM6  
DT 01-OCT-2000 (TREMELREL. 15, Created)  
DT 01-OCT-2000 (TREMELREL. 15, Last sequence update)  
DT 01-JUN-2002 (TREMELREL. 21, Last annotation update)  
DE Topoisomerase IV subunit.  
GN XF1353.  
OS Xylella fastidiosa.  
OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;  
OC Xylella.  
OC NCBI\_TaxID=2371;  
RN [1]  
RN SEQUENCE FROM N.A.  
RC STRAIN=9A5C;  
RX MEDLINE=20365717; PubMed=10910347;  
RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,  
RA Alvares R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,  
RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,  
RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carrer H.,  
RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,  
RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,  
RA Pacinani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,  
RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,  
RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,  
RA Ho P.L., Hoheisel J.D., Jungueira M.L., Kemper E.L., Kitajima J.P.,  
RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,  
RA Lemos E.G.M., Lemos M.V.F., Lopes F., Lopes C.R., Machado J.A.,  
RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,  
RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,  
RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,  
RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,  
RA Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,  
RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,  
RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,  
RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,  
RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,  
RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,  
RA da Silveira J.F., Silvestri M.L.Z., Silveira W.J., de Souza A.A.,  
RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tshako M.H.,  
RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,  
RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;  
RT "The genome sequence of the plant pathogen Xylella fastidiosa.";  
RL Nature 406:151-159(2000).  
DR EMBL; AE003967; AAF84162.1; .  
DR HSP; P09097; 1AB4.  
DR InterPro: IPR002205; DNA\_topoisomIV.  
DR Pfam; PF00521; DNA\_topoisomIV; 1.  
DR ProDom; PD000742; DNA\_topoisomIV; 1.  
DR SMART; SM00434; TOP4c; 1.  
DR TIGRFAMs; TIGR01062; parC\_Gneg; 1.  
KW Complete proteome.  
SQ SEQUENCE 749 AA; 83344 MW; 4DCD10F480EE0257 CRC64;  
Query Match 67.3%; Score 35; DB 16; Length 749;  
Best Local Similarity 77.8%; Pred. No. 88;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 2 EVVPXGMSY 10  
DB 526 EVDPSGMSY 534  
RESULT 14  
Q14122  
ID Q14122 PRELIMINARY; PRT; 1902 AA.  
AC Q14122;  
DT 01-NOV-1996 (TREMELREL. 01, Created)  
DT 01-NOV-1996 (TREMELREL. 01, Last sequence update)  
DT 01-NOV-2002 (TREMELREL. 20, Last annotation update)  
DE DNA-binding protein (Mbp-1) (Fragment).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RN SEQUENCE FROM N.A.  
RX MEDLINE=90205817; PubMed=2108316;  
RA Baldwin A.S., LeClair K.P., Singh H., Sharp P.A.;  
RT "A large protein containing zinc finger domains binds to related  
RT sequence elements in the enhancers of the class I major  
RT histocompatibility complex and kappa immunoglobulin genes.";  
RL Mol. Cell. Biol. 10:1406-1414(1990).  
DR EMBL; M32019; AAA17534.1; .  
DR HSP; P15822; 1BBO.  
DR InterPro: IPR000822; Znf\_C2H2.  
DR Pfam; PF00096; zf-C2H2; 3.  
DR SMART; SM00355; Znf\_C2H2; 2.  
DR PROSITE; PS00028; ZINC\_FINGER\_C2H2\_1; 2.  
DR PROSITE; PS0157; ZINC\_FINGER\_C2H2\_2; 2.  
KW DNA-binding; Metal-binding; Zinc-finger.  
FT NON\_TER 1  
SQ SEQUENCE 1902 AA; 207457 MW; 35BBF0D961E6ED20 CRC64;  
Query Match 67.3%; Score 35; DB 4; Length 1902;  
Best Local Similarity 66.7%; Pred. No. 2.4e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
QY 3 VVPXGMSYS 11  
DB 1590 VVPAGLTYS 1598  
RESULT 15  
Q12479  
ID Q12479 PRELIMINARY; PRT; 156 AA.  
AC Q12479;  
DT 01-NOV-1996 (TREMELREL. 01, Created)  
DT 01-NOV-1996 (TREMELREL. 01, Last sequence update)  
DT 01-JUN-2002 (TREMELREL. 21, Last annotation update)  
DE ORF YOR013W.  
GN YOR013W.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
OX NCBI\_TaxID=4932;  
RN [1]  
RN SEQUENCE FROM N.A.  
RA De Haan M., Grivell L.A., Maarse A.C.;  
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RN MIPS;  
RN SEQUENCE FROM N.A.  
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RN SEQUENCE FROM N.A.  
RC STRAIN=FY1679;  
RA De Haan M., Maarse A.C., Grivell L.A.;  
RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RN SEQUENCE FROM N.A.  
RC STRAIN=FY1679;  
RX MEDLINE=94019318; PubMed=8413243;  
RA Dumont M.E., Schlachter J.B., Cardillo T.S., Hayes M.K., Bethlenny G.,  
RA Sherman F.;  
RT "CYC2 encodes a factor involved in mitochondrial import of yeast  
RT cytochrome c.";  
RL Mol. Cell. Biol. 13:6442-6451(1993).  
RN [5]  
RN SEQUENCE FROM N.A.  
RC STRAIN=FY1679;  
RX MEDLINE=94169519; PubMed=7764548;  
RA Lee Y.S., Shimizu J., Yoda K., Yamasaki M.;  
RT "Molecular cloning of a gene, DHS1, which complements a drug-  
RT hypersensitive mutation of the yeast Saccharomyces cerevisiae.";  
RL Biosci. Biotechnol. Biochem. 58:391-395(1994).

```

DR EMBL; Z74920; CAA99201.1; -.
DR EMBL; X87331; CAA60762.1; -.
DR SGD; S0005539; YOR013W.
SQ SEQUENCE 156 AA; 17881 MW; 380442B74C272B41 CRC64;

Query Match 65.4%; Score 34; DB 3; Length 156;
Best Local Similarity 66.7%; Pred. No. 27;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 EVVPXGMSY 10
Db ||| |||
50 EVWPLGMDY 58

RESULT 16
Q971S2 PRELIMINARY; PRT; 219 AA.
AC Q971S2;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DE Putative ribose 5-phosphate isomerase.
DS sri302.
OS Sulfolobus tokodaii.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.
OX NCBI_TaxID=111955;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JCM 10545 / 7;
RX PubMed=11572479;
RA Kawarabayashi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,
RA Sekine M., Baba S.-I., Ankai A., Kosugi H., Hosoyama A., Fukui S.,
RA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,
RA Yoshizawa T., Tanaka T., Kudoh Y., Yamazaki J., Kishida N., Oguchi A.,
RA Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A.,
RA Oshima T., Kikuchi H.;
RT "Complete genome sequence of an aerobic thermoacidophilic
RT Crenarchaeon, Sulfolobus tokodaii strain7."
RL DNA Res. 8:123-140(2001).
DR EMBL; AP000985; BAB66348.1; -.
DR InterPro; IPR004788; RpiA.
DR ProDom; PD005813; RpiA; 1.
KW Isomerase; Hypothetical protein; Complete proteome.
SQ SEQUENCE 219 AA; 24541 MW; A4E9A3F2C4006D90 CRC64;

Query Match 65.4%; Score 34; DB 17; Length 219;
Best Local Similarity 66.7%; Pred. No. 38;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 EVVPXGMSY 10
Db ||| |||
131 EVVPGVAY 139

RESULT 17
Q28342 PRELIMINARY; PRT; 252 AA.
ID Q28342;
AC Q28342;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Cell division inhibitor (MIND-2).
DS AF1937.
OS Archaeoglobus fulgidus.
OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
OC Archaeoglobaceae; Archaeoglobus.
OX NCBI_TaxID=2234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=98049343; PubMed=9389475;
RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,

```

```

RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
RA Richardson D.L., Kerlavage A.R., Graham D.E., Kyripides N.C.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
RA Cotton M.D., Spriggs T., Artiach P., Kaine B.P., Sykes S.M.,
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RT "The complete genome sequence of the hyperthermophilic, sulphate-
RT reducing archaeon Archaeoglobus fulgidus.";
RL Nature 390:364-370(1997).
DR EMBL; AE000970; AAB89318.1; -.
DR TIGR; AF1937; -.
DR InterPro; IPR000707; ATPase_Para.
DR Pfam; PF00991; Para; 1.
KW Hypothetical protein; Cell division; Complete proteome.
SQ SEQUENCE 252 AA; 27130 MW; A401DC1F93E8C538 CRC64;

Query Match 65.4%; Score 34; DB 17; Length 252;
Best Local Similarity 75.0%; Pred. No. 44;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 EVVPXGMS 9
Db ||| |||
81 EVIPAGMS 88

RESULT 18
Q96MU1 PRELIMINARY; PRT; 290 AA.
ID Q96MU1;
AC Q96MU1;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DS CDNA FL331891 fis, clone NT2RP7003304, weakly similar to YceA protein
DE homolog ybfQ.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Tashiro H., Yamazaki M., Watanabe K., Kumagai A., Itakura S.,
RA Fukuzumi Y., Fujimori Y., Komiyama M., Sugiyama T., Irie R.,
RA Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,
RA Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K., Wagatsuma M.,
RA Murakawa K., Kanehori K., Takahashi-Fujii A., Oshima A., Sugiyama A.,
RA Kawakami B., Suzuki Y., Sugano S., Nagahari K., Masuho Y., Nagai K.,
RA Isogai T.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK056453; BAB71188.1; -.
DR InterPro; IPR001763; Rhodanese-like.
DR Pfam; PF00581; Rhodanese; 1.
SQ SEQUENCE 290 AA; 32972 MW; 41FB7FB0217C1421 CRC64;

Query Match 65.4%; Score 34; DB 4; Length 290;
Best Local Similarity 66.7%; Pred. No. 52;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXGMS 9
Db ||| |||
35 EEIVPMGIS 43

RESULT 19
Q92MD6 PRELIMINARY; PRT; 387 AA.
ID Q92MD6
AC Q92MD6;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)

```

```

DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DE 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE Putative HIPURATE hydrolase protein (EC 3.5.1.32).
GN HIPOL OR R02690 OR SMC00682.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RX MEDLINE=21396507; PubMed=11481430;
RA Capela D., Barloy-Hubier F., Gouzy J., Bothe G., Ampe F., Batut J.,
RA Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
RA Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masuy D.,
RA Pohl T., Portetelle D., Puehler A., Purnelle B., Ramsperger U.,
RA Renard C., Thebault P., Vandenbol M., Weidner S., Galibert F.;
RT "Analysis of the chromosome sequence of the legume symbiont
RT Sinorhizobium meliloti strain 1021."
RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
DR EMBL; AL591791; CAC47269.1; -.
DR InterPro; IPR002933; Peptidase_M20.
DR Pfam; PF01546; Peptidase_M20; 1.
KW Hydrolase; Complete proteome.
SQ SEQUENCE 387 AA; 41074 MW; 342763088907A6E3 CRC64;

Query Match 65.4%; Score 34; DB 16; Length 387;
Best Local Similarity 50.0%; Pred. No. 70;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10
: | | | |
DB 367 DEAIPIHGISY 376

RESULT 20
Q98BP5 ID Q98BP5 PRELIMINARY; PRT; 541 AA.
AC Q98BP5;
DT 01-OCT-2001 (TReMBLrel. 18, Created)
DT 01-OCT-2001 (TReMBLrel. 18, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Probable DNA ligase.
GN MLL5481.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Idegawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpou S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti."
RL DNA Res. 7:331-338(2000).
DR EMBL; AP003006; BAB51927.1; -.
DR InterPro; IPR000977; DNA_ligase.
DR Pfam; PF01068; DNA_ligase; 1.
DR PROSITE; PS00697; DNA_LIGASE_A1; UNKNOWN_1.
DR PROSITE; PS0160; DNA_LIGASE_A3; 1.
KW Ligase; Complete proteome.
SQ SEQUENCE 541 AA; 60645 MW; 2EEFF705453F28F8 CRC64;

Query Match 65.4%; Score 34; DB 16; Length 541;
Best Local Similarity 60.0%; Pred. No. 1e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10
: | | | |
DB 367 DEAIPIHGISY 376

```

```

DB 445 EELVPVGKAY 454
: | | | | | | |
: | | | | | | |

RESULT 21
Q9PDD2 ID Q9PDD2 PRELIMINARY; PRT; 544 AA.
AC Q9PDD2;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE Putative ABC substrate-binding protein-iron.
GN ABCSBP-5 OR U0359.
OS Ureaplasma parvum (Ureaplasma urealyticum biotype 1).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Ureaplasma.
OX NCBI_TaxID=134821;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SEROVAR 3;
RX MEDLINE=20500219; PubMed=11048724;
RA Glass J.I., Lefkowitz E.J., Glass J.S., Heiner C.R., Chen E.Y.,
RA Cassell G.H.;
RT "The complete sequence of the mucosal pathogen Ureaplasma
RT urealyticum."
RL Nature 407:757-762(2000).
DR EMBL; AE002133; AAF30768.1; -.
KW Complete proteome.
SQ SEQUENCE 544 AA; 61291 MW; CF8756202A389C00 CRC64;

Query Match 65.4%; Score 34; DB 16; Length 544;
Best Local Similarity 70.0%; Pred. No. 1e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10
: | | | | |
DB 135 EEVPHYLSY 144

RESULT 22
Q9URR4 ID Q9URR4 PRELIMINARY; PRT; 842 AA.
AC Q9URR4;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Sulfate permease SufB.
GN SUTB.
OS Penicillium chrysogenum.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Penicillium.
OX NCBI_TaxID=5076;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Q176;
RX MEDLINE=20042342; PubMed=10572125;
RA Van de Kamp M., Pizzini E., Vos A., Van der Lende T.R.,
RA Schuur T.A., Newbert R.W., Turner G., Konings W.N., Driessen A.J.M.;
RT "Sulfate Transport in Penicillium chrysogenum: Cloning and
RT Characterization of the suta and sutB Genes."
RL J. Bacteriol. 181:7228-7234(1999).
DR EMBL; AF163974; AAF14539.1; -.
DR InterPro; IPR002645; STAS.
DR InterPro; IPR001902; Sulfate_transp.
DR Pfam; PF01740; STAS; 1.
DR Pfam; PF00916; Sulfate_transp; 1.
DR TIGRfams; TIGR00815; sulp; 1.
DR PROSITE; PS01130; SULFATE_TRANSP; UNKNOWN_1.
SQ SEQUENCE 842 AA; 91865 MW; 839A55486E733D15 CRC64;

Query Match 65.4%; Score 34; DB 3; Length 842;
Best Local Similarity 66.7%; Pred. No. 1.6e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```

QY 3 VVPXGMSYS 11  
 DB 111 VPOGMAYA 119

## RESULT 23

Q8XT05 Q8XT05 PRELIMINARY; PRT; 1049 AA.  
 AC Q8XT05;  
 DT 01-MAR-2002 (TREMBlrel. 20, Created)  
 DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)  
 DE 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE Probable multidrug efflux system transmembrane protein.  
 GN MEXD OR RSP0312 OR RS03457.  
 OS Ralstonia solanacearum (Pseudomonas solanacearum).  
 OG Plasmid megaplasmid.  
 OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;  
 OC Ralstonia.  
 OX NCBI\_TaxID=305;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX STRAIN=GM11000;  
 RC MEDLINE=21681879; PubMed=11823852;  
 RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,  
 RA Arlat M., Billault A., Brottier P., Camus J.C., Cattolico L.,  
 RA Chandler M., Choise N., Claudel-Renard C., Cunnac S., Demange N.,  
 RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schiex T.,  
 RA Siguer P., Thebaud P., Whalen M., Wincker P., Levy M.,  
 RA Welssbach J., Boucher C.A.;  
 RT "Genome sequence of the plant pathogen Ralstonia solanacearum."  
 RL Nature 415:497-502(2002).  
 DR EMBL; AL646078; CAD17463.1; -.  
 DR InterPro; IPR001036; Acriflavin\_res.  
 DR InterPro; IPR004764; HAE1.  
 DR Pfam; PF00873; ACR\_tran.1.  
 DR PRINTS; PR00702; ACRIFLAVINRP.  
 DR TIGRFS; TIGR00915; 2A0602; 1.  
 DR PROSITE; PSS0156; SSD; 1.  
 KW Plasmid; Complete proteome.  
 SQ SEQUENCE 1049 AA; 111769 MW; CB59674B670089CE CRC64;

Query Match 65.4%; Score 34; DB 16; Length 1049;  
 Best Local Similarity 75.0%; Pred. No. 2.1e+02;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 VVPXGMSYS 11  
 DB 317 MPAGMSYS 324

## RESULT 24

Q96914 Q96914 PRELIMINARY; PRT; 1499 AA.  
 AC Q96914;  
 DT 01-DEC-2001 (TREMBlrel. 19, Created)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
 DE Putative aminophospholipid translocase (Aminophospholipid-transporting ATPase).  
 GN ATP10C.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=2125279; PubMed=11326269;  
 RA Meguro M., Kashiwagi A., Mitsuya K., Nakao M., Kondo I., Saitoh S.,  
 RA Oshimura M.;  
 RT "A novel maternally expressed gene, ATP10C, encodes a putative aminophospholipid translocase associated with Angelman syndrome.";

RL Nat. Genet. 28:19-20(2001).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21313119; PubMed=11353404;  
 RA Herzling L.B.K., Kim S.J., Cook E.H. Jr., Ledbetter D.H.;  
 RT "The human aminophospholipid-transporting ATPase gene ATP10C maps adjacent to UBE3A and exhibits similar imprinted expression.";  
 RT Am. J. Hum. Genet. 68:1501-1505(2001).  
 DR EMBL; AB051358; BAB47392.1; -.  
 DR EMBL; AY029504; AAK33100.1; JOINED.  
 DR EMBL; AY029487; AAK33100.1; JOINED.  
 DR EMBL; AY029488; AAK33100.1; JOINED.  
 DR EMBL; AY029489; AAK33100.1; JOINED.  
 DR EMBL; AY029490; AAK33100.1; JOINED.  
 DR EMBL; AY029491; AAK33100.1; JOINED.  
 DR EMBL; AY029492; AAK33100.1; JOINED.  
 DR EMBL; AY029493; AAK33100.1; JOINED.  
 DR EMBL; AY029494; AAK33100.1; JOINED.  
 DR EMBL; AY029495; AAK33100.1; JOINED.  
 DR EMBL; AY029496; AAK33100.1; JOINED.  
 DR EMBL; AY029497; AAK33100.1; JOINED.  
 DR EMBL; AY029498; AAK33100.1; JOINED.  
 DR EMBL; AY029499; AAK33100.1; JOINED.  
 DR EMBL; AY029500; AAK33100.1; JOINED.  
 DR EMBL; AY029501; AAK33100.1; JOINED.  
 DR EMBL; AY029502; AAK33100.1; JOINED.  
 DR EMBL; AY029503; AAK33100.1; JOINED.  
 DR InterPro; IPR001757; ATPase\_E1-E2.  
 DR InterPro; IPR001064; Crystallin.  
 DR Pfam; PF00702; Hydrolase; 1.  
 DR PROSITE; PS00154; ATPASE\_E1\_E2; UNKNOWN\_1.  
 DR PROSITE; PS00225; CRYSTALLIN\_BETAGAMMA; UNKNOWN\_1.  
 SQ SEQUENCE 1499 AA; 167687 MW; D4996A4D0635A68D CRC64;

Query Match 65.4%; Score 34; DB 4; Length 1499;  
 Best Local Similarity 72.7%; Pred. No. 3e+02;  
 Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSYS 11  
 DB 469 EEVPRGGSVS 479

## RESULT 25

Q8TX62 Q8TX62 PRELIMINARY; PRT; 143 AA.  
 AC Q8TX62;  
 DT 01-JUN-2002 (TREMBlrel. 21, Created)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE Uncharacterized conserved protein.  
 GN MK0814.  
 OS Methanopyrus kandleri.  
 OC Archaea; Euryarchaeota; Methanopyri; Methanopyrales; Methanopyraceae;  
 OC Methanopyrus  
 OX NCBI\_TaxID=2320;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX STRAIN=AV19 / DSM 6324 / JCM 9639;  
 RA MEDLINE=21927647; PubMed=11930014;  
 RA Slesarev A.I., Mezheva K.V., Makarova K.S., Polushin N.N.,  
 RA Shcherbinina O.V., Shakhova V.V., Belova G.I., Aravind L.,  
 RA Natile D.A., Rogozin I.B., Tatusov R.L., Wolf Y.I., Stetter K.O.,  
 RA Malykh A.G., Koonin E.V., Kozyavkin S.A.;  
 RT "The complete genome of hyperthermophile Methanopyrus kandleri AV19 and monophyly of archaeal methanogens."  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:4644-4649(2002).  
 DR EMBL; AE010372; AAM02027.1; -.  
 KW Complete proteome.  
 SQ SEQUENCE 143 AA; 15734 MW; 4C8B28A1FBEDD0B CRC64;

Query Match 63.5%; Score 33; DB 17; Length 143;

Best Local Similarity 60.0%; Pred. No. 40;  
Matches 6; Conservative 1; Mismatches

3; Indels 0; Gaps 0;

Qy 1 EEVVPXGMSY 10

||:| | |

Db 75 EELVPQGAGY 84

Search completed: June 10, 2003, 13:46:26  
Job time : 29.7857 secs



GenCore version 5.1.1.6

Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2003, 13:24:45 ; Search time 31.3571 Seconds

(without alignments)

46.744 Million cell updates/sec

Title: US-09-909-164-6

Perfect score: 52

Sequence: 1 ERVFXGMSYS 11

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 75 summaries

Database :

A\_Geneseq\_101002.\*

1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
 2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
 3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
 4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
 5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
 6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
 7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
 8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
 9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
 10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
 11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
 12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
 13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
 14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
 15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
 16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
 17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
 18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
 19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
 20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
 21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
 22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
 23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	50	96.2	11	23	ABB80521
2	50	96.2	11	23	ABB80522
3	50	96.2	11	23	ABB80525
4	50	96.2	11	23	ABB80526
5	50	96.2	11	23	ABB80559
6	50	96.2	11	23	ABB80563
7	50	96.2	11	23	ABB80564
8	50	96.2	11	23	ABB80565
9	50	96.2	11	23	ABB80566
10	50	96.2	11	23	ABB80567

11	50	96.2	11	23	ABB80568	Hepatitis C virus
12	46	88.5	11	23	ABB80524	Hepatitis C virus
13	46	88.5	11	23	ABB80528	Hepatitis C virus
14	46	88.5	11	23	ABB80529	Hepatitis C virus
15	46	88.5	11	23	ABB80561	Hepatitis C virus
16	46	88.5	11	23	ABB80562	Hepatitis C virus
17	45	86.5	11	23	ABB80523	Hepatitis C virus
18	45	86.5	11	23	ABB80527	Hepatitis C virus
19	45	86.5	11	23	ABB80535	Hepatitis C virus
20	45	86.5	11	23	ABB80536	Hepatitis C virus
21	45	86.5	11	23	ABB80539	Hepatitis C virus
22	45	86.5	11	23	ABB80540	Hepatitis C virus
23	45	86.5	11	23	ABB80558	Hepatitis C virus
24	45	86.5	11	23	ABB80560	Hepatitis C virus
25	44	84.6	11	23	ABB80544	Hepatitis C virus
26	44	84.6	11	23	ABB80545	Hepatitis C virus
27	44	84.6	11	23	ABB80549	Hepatitis C virus
28	44	84.6	11	23	ABB80552	Hepatitis C virus
29	44	84.6	11	23	ABB80553	Hepatitis C virus
30	42	80.8	11	23	ABB80530	Hepatitis C virus
31	41	78.8	11	23	ABB80538	Hepatitis C virus
32	41	78.8	11	23	ABB80542	Hepatitis C virus
33	41	78.8	11	23	ABB80543	Hepatitis C virus
34	40	76.9	11	23	ABB80537	Hepatitis C virus
35	40	76.9	11	23	ABB80541	Hepatitis C virus
36	40	76.9	11	23	ABB80547	Hepatitis C virus
37	40	76.9	11	23	ABB80548	Hepatitis C virus
38	40	76.9	11	23	ABB80551	Hepatitis C virus
39	40	76.9	11	23	ABB80556	Hepatitis C virus
40	40	76.9	11	23	ABB80557	Hepatitis C virus
41	40	76.9	20	20	AAU76810	Hepatitis C virus
42	40	76.9	1022	22	ABG03621	Novel human diago
43	40	76.9	1022	22	ABG05826	Novel human diago
44	40	76.9	1022	22	ABG08173	Hepatitis C virus
45	39	75.0	11	23	ABB80546	Hepatitis C virus
46	39	75.0	11	23	ABB80550	Hepatitis C virus
47	39	75.0	11	23	ABB80554	Hepatitis C virus
48	39	75.0	11	23	ABB80555	Hepatitis C virus
49	38	73.1	11	23	ABB80533	Hepatitis C virus
50	38	73.1	11	23	ABB80534	Hepatitis C virus
51	38	73.1	3472	21	AAV90913	Cenarchaeum symbio
52	37	71.2	11	23	ABB80531	Hepatitis C virus
53	37	71.2	11	23	ABB80532	Hepatitis C virus
54	36	69.2	244	21	AA812881	Murine JNK3 bindin
55	36	69.2	484	21	AA812882	Murine JNK3 bindin
56	35	67.3	11	18	AAW99288	Peptide N424 from
57	34	65.4	842	21	AAV44359	P. chrysogenum sut
58	34	65.4	947	21	AA825105	Human novel protei
59	34	65.4	1070	22	AAU14378	Human novel protei
60	33	63.5	12	21	AAV83772	HCV NS3 protease s
61	33	63.5	12	21	AAV83774	HCV NS3 protease s
62	33	63.5	13	18	AAW99276	Peptide D4 from WO
63	33	63.5	13	18	AAW99275	Peptide 5 used in
64	33	63.5	14	18	AAW99275	Peptide D3 from WO
65	33	63.5	14	18	AAW99277	Peptide C0 from WO
66	33	63.5	14	18	AAW33285	Peptide 4 used in
67	33	63.5	14	18	AAW33287	Peptide 6 used in
68	33	63.5	16	18	AAW99274	Peptide D2 from WO
69	33	63.5	16	18	AAW33284	Peptide 3 used in
70	33	63.5	18	18	AAW99273	Peptide D1 from WO
71	33	63.5	18	18	AAW33283	Peptide 2 used in
72	33	63.5	18	19	AAW71281	Cleavable substrate
73	33	63.5	20	18	AAW99272	Peptide PS from WO
74	33	63.5	20	18	AAW33282	Peptide 1 used in
75	33	63.5	20	20	AAV14511	HCV NS3 protease t

## ALIGNMENTS

RESULT 1  
 ABB80521  
 ID ABB80521 standard; peptide; 11 AA.





FT Misc-difference 8 /note= "D-form residue"  
FT Modified-site 11  
FT /note= "C-terminal amide"  
XX  
PN WO200208251-A2.  
XX  
XX 31-JAN-2002.  
XX  
XX 19-JUL-2001; 2001WO-US23169.  
XX  
XX 21-JUL-2000; 2000US-220101P.  
XX  
XX (CORV-) CORVAS INT INC.  
XX  
XX Llm-wilby M, Levy OE, Brunck TK;  
XX WPI; 2002-361643/39.  
XX  
XX Novel peptide compound having hepatitis C virus protease inhibitory  
PT activity useful for treating disorders associated with hepatitis C  
PT virus protease  
XX  
XX Claim 17; Page 64; 69pp; English.  
XX  
XX The sequence represents a peptide compound of the invention having  
CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
CC invention are alpha-ketoamide peptide analogues. The peptides have  
CC virucide activity, and are useful for treating and in the manufacture of  
CC a medicament to treat disorders associated with HCV protease. A  
CC pharmaceutical composition comprising the peptide as an active ingredient  
CC is useful for treating disorders associated with hepatitis C virus.  
XX  
SQ Sequence 11 AA;  
Query Match 96.2%; Score 50; DB 23; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.0011;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 EEVVPXGMSYS 11  
Db 1 EEVVPXGMSYS 11  
RESULT 4  
ABB80526  
ID ABB80526 standard; peptide; 11 AA.  
AC ABB80526;  
XX  
XX 08-OCT-2002 (first entry)  
XX  
XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #6.  
DE Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
KW virucide.  
XX  
XX Synthetic.  
XX  
XX Key Location/Qualifiers  
FH Modified-site 1 /note= "N-terminal acetyl"  
FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with  
FT residue 7"  
FT Misc-difference 8 /note= "D-form residue"  
FT Misc-difference 9 /note= "D-form residue"  
FT Modified-site 11 /note= "C-terminal amide"  
XX  
XX WO200208251-A2.

FT Misc-difference 8 /note= "D-form residue"  
FT Modified-site 11  
FT /note= "C-terminal amide"  
XX  
PN WO200208251-A2.  
XX  
XX 31-JAN-2002.  
XX  
XX 19-JUL-2001; 2001WO-US23169.  
XX  
XX 21-JUL-2000; 2000US-220101P.  
XX  
XX (CORV-) CORVAS INT INC.  
XX  
XX Llm-wilby M, Levy OE, Brunck TK;  
XX WPI; 2002-361643/39.  
XX  
XX Novel peptide compound having hepatitis C virus protease inhibitory  
PT activity useful for treating disorders associated with hepatitis C  
PT virus protease  
XX  
XX Claim 17; Page 64; 69pp; English.  
XX  
XX The sequence represents a peptide compound of the invention having  
CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
CC invention are alpha-ketoamide peptide analogues. The peptides have  
CC virucide activity, and are useful for treating and in the manufacture of  
CC a medicament to treat disorders associated with HCV protease. A  
CC pharmaceutical composition comprising the peptide as an active ingredient  
CC is useful for treating disorders associated with hepatitis C virus.  
XX  
SQ Sequence 11 AA;  
Query Match 96.2%; Score 50; DB 23; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.0011;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 EEVVPXGMSYS 11  
Db 1 EEVVPXGMSYS 11  
RESULT 5  
ABB80559  
ID ABB80559 standard; peptide; 11 AA.  
AC ABB80559;  
XX  
XX 08-OCT-2002 (first entry)  
XX  
XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #39.  
DE Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
KW virucide.  
XX  
XX Synthetic.  
XX  
XX Key Location/Qualifiers  
FH Modified-site 1 /note= "N-terminal acetyl"  
FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with  
FT residue 7"  
FT Misc-difference 8 /note= "D-form residue"  
FT Modified-site 8 /note= "Oxymethionine"  
FT Modified-site 11 /note= "C-terminal amide"  
XX  
XX WO200208251-A2.  
XX  
XX 31-JAN-2002.  
XX  
XX 19-JUL-2001; 2001WO-US23169.  
XX  
XX 21-JUL-2000; 2000US-220101P.

XX 31-JAN-2002.  
XX  
XX 19-JUL-2001; 2001WO-US23169.  
XX  
XX 21-JUL-2000; 2000US-220101P.  
XX  
XX (CORV-) CORVAS INT INC.  
XX  
XX Llm-wilby M, Levy OE, Brunck TK;  
XX WPI; 2002-361643/39.  
XX  
XX Novel peptide compound having hepatitis C virus protease inhibitory  
PT activity useful for treating disorders associated with hepatitis C  
PT virus protease  
XX  
XX Claim 17; Page 64; 69pp; English.  
XX  
XX The sequence represents a peptide compound of the invention having  
CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
CC invention are alpha-ketoamide peptide analogues. The peptides have  
CC virucide activity, and are useful for treating and in the manufacture of  
CC a medicament to treat disorders associated with HCV protease. A  
CC pharmaceutical composition comprising the peptide as an active ingredient  
CC is useful for treating disorders associated with hepatitis C virus.  
XX  
SQ Sequence 11 AA;  
Query Match 96.2%; Score 50; DB 23; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.0011;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 EEVVPXGMSYS 11  
Db 1 EEVVPXGMSYS 11  
RESULT 5  
ABB80559  
ID ABB80559 standard; peptide; 11 AA.  
AC ABB80559;  
XX  
XX 08-OCT-2002 (first entry)  
XX  
XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #39.  
DE Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
KW virucide.  
XX  
XX Synthetic.  
XX  
XX Key Location/Qualifiers  
FH Modified-site 1 /note= "N-terminal acetyl"  
FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with  
FT residue 7"  
FT Misc-difference 8 /note= "D-form residue"  
FT Modified-site 8 /note= "Oxymethionine"  
FT Modified-site 11 /note= "C-terminal amide"  
XX  
XX WO200208251-A2.  
XX  
XX 31-JAN-2002.  
XX  
XX 19-JUL-2001; 2001WO-US23169.  
XX  
XX 21-JUL-2000; 2000US-220101P.

XX PA (CORV-) CORVAS INT INC.  
XX PI Lim-wilby M, Levy OE, Brunck TK;  
XX DR WPI; 2002-361643/39.  
XX PT Novel peptide compound having hepatitis C virus protease inhibitory  
XX PT activity useful for treating disorders associated with hepatitis C  
XX PT virus protease.  
XX PS  
XX SQ Claim 17; Page 65; 69pp; English.  
XX  
CC The sequence represents a peptide compound of the invention having  
CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
CC invention are alpha-ketoamide peptide analogues. The peptides have  
CC virucide activity, and are useful for treating and in the manufacture of  
CC a medicament to treat disorders associated with HCV protease. A  
CC pharmaceutical composition comprising the peptide as an active ingredient  
CC is useful for treating disorders associated with hepatitis C virus.  
XX SQ Sequence 11 AA;

Query Match 96.2%; Score 50; DB 23; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.0011;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEVVPXGMSYS 11  
DB 1 EEVVPXGMSYS 11

RESULT 6  
ABB0563  
ID ABB0563 standard; peptide; 11 AA.  
XX  
AC ABB0563;  
XX  
DT 08-OCT-2002 (first entry)  
XX  
DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #43.  
XX  
KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
KW virucide.  
XX  
OS Synthetic.

XX FH Key Location/Qualifiers  
XX FT Modified-site 1 /note= "N-terminal acetyl"  
XX FT Modified-site 6 /note= "Valyl carbonyl forming keto-amide linkage with  
XX FT residue 7"  
XX FT Modified-site 11 /note= "C-terminal amide"

XX PN WO200208251-A2.  
XX PD 31-JAN-2002.  
XX PF 19-JUL-2001; 2001WO-US23169.  
XX PR 21-JUL-2000; 2000US-220101P.  
XX PA (CORV-) CORVAS INT INC.

XX PI Lim-wilby M, Levy OE, Brunck TK;  
XX DR WPI; 2002-361643/39.  
XX PT Novel peptide compound having hepatitis C virus protease inhibitory  
XX PT activity useful for treating disorders associated with hepatitis C  
XX PT virus protease.

XX PS Claim 17; Page 65; 69pp; English.  
XX

CC The sequence represents a peptide compound of the invention having  
CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
CC invention are alpha-ketoamide peptide analogues. The peptides have  
CC virucide activity, and are useful for treating and in the manufacture of  
CC a medicament to treat disorders associated with HCV protease. A  
CC pharmaceutical composition comprising the peptide as an active ingredient  
CC is useful for treating disorders associated with hepatitis C virus.  
XX SQ Sequence 11 AA;

Query Match 96.2%; Score 50; DB 23; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.0011;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEVVPXGMSYS 11  
DB 1 EEVVPXGMSYS 11

RESULT 7  
ABB0564  
ID ABB0564 standard; peptide; 11 AA.  
XX  
AC ABB0564;  
XX  
DT 08-OCT-2002 (first entry)  
XX  
DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #44.  
XX  
KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
KW virucide.  
XX  
OS Synthetic.

XX FH Key Location/Qualifiers  
XX FT Modified-site 1 /note= "N-terminal acetyl"  
XX FT Modified-site 6 /note= "Leucyl carbonyl forming keto-amide linkage with  
XX FT residue 7"  
XX FT Modified-site 11 /note= "C-terminal amide"

XX PN WO200208251-A2.  
XX PD 31-JAN-2002.  
XX PF 19-JUL-2001; 2001WO-US23169.  
XX PR 21-JUL-2000; 2000US-220101P.

XX PA (CORV-) CORVAS INT INC.  
XX PI Lim-wilby M, Levy OE, Brunck TK;  
XX DR WPI; 2002-361643/39.

XX Novel peptide compound having hepatitis C virus protease inhibitory  
XX activity useful for treating disorders associated with hepatitis C  
XX virus protease.

XX PS Claim 17; Page 65; 69pp; English.

XX The sequence represents a peptide compound of the invention having  
XX hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
XX invention are alpha-ketoamide peptide analogues. The peptides have  
XX virucide activity, and are useful for treating and in the manufacture of  
XX a medicament to treat disorders associated with HCV protease. A  
XX pharmaceutical composition comprising the peptide as an active ingredient  
XX is useful for treating disorders associated with hepatitis C virus.

XX SQ Sequence 11 AA;

Query Match 96.2%; Score 50; DB 23; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.0011;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEVVPXGMSYS 11  
|||||  
Db 1 EEVVPXGMSYS 11

## RESULT 8

ABB80565  
ID ABB80565 standard; peptide; 11 AA.

XX AC ABB80565;

XX DT 08-OCT-2002 (first entry)

XX DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #45.

XX KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
XX KW virucide.

XX OS Synthetic.

XX FH Key Location/Qualifiers

FT Modified-site 1 /note= "N-terminal acetyl"

FT Modified-site 6

FT /note= "Morleucyl carbonyl forming keto-amide linkage  
with residue 7"

FT Modified-site 11

FT /note= "C-terminal amide"

XX WO200208251-A2.

XX 31-JAN-2002.

XX 19-JUL-2001; 2001WO-US23169.

XX 21-JUL-2000; 2000US-220101P.

XX (CORV-) CORVAS INT INC.

XX Lim-wilby M, Levy OE, Brunck TK;

XX WPI; 2002-361643/39.

XX Novel peptide compound having hepatitis C virus protease inhibitory  
PT activity useful for treating disorders associated with hepatitis C  
PT virus protease

XX Claim 17; Page 65; 69pp; English.

XX The sequence represents a peptide compound of the invention having  
CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
CC invention are alpha-ketoamide peptide analogues. The peptides have  
CC virucide activity, and are useful for treating and in the manufacture of  
CC a medicament to treat disorders associated with HCV protease. A  
CC pharmaceutical composition comprising the peptide as an active ingredient  
CC is useful for treating disorders associated with hepatitis C virus.

XX SQ Sequence 11 AA;

Query Match 96.2%; Score 50; DB 23; Length 11;

Best Local Similarity 100.0%; Pred. No. 0.0011;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEVVPXGMSYS 11  
|||||  
Db 1 EEVVPXGMSYS 11

## RESULT 9

ABB80566

ID ABB80566 standard; peptide; 11 AA.

XX AC ABB80566;

XX DT 08-OCT-2002 (first entry)

XX DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #46.

XX KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
XX KW virucide.

XX OS Synthetic.

XX FH Key Location/Qualifiers

FT Modified-site 1 /note= "N-terminal acetyl"

FT Modified-site 6

FT /note= "2-aminoisobutyl carbonyl residue forming a  
keto-amide linkage with residue 7"

FT Modified-site 11

FT /note= "C-terminal amide"

XX WO200208251-A2.

XX 31-JAN-2002.

XX 19-JUL-2001; 2001WO-US23169.

XX 21-JUL-2000; 2000US-220101P.

XX (CORV-) CORVAS INT INC.

XX Lim-wilby M, Levy OE, Brunck TK;

XX WPI; 2002-361643/39.

XX Novel peptide compound having hepatitis C virus protease inhibitory  
PT activity useful for treating disorders associated with hepatitis C  
PT virus protease

XX Claim 17; Page 65; 69pp; English.

XX The sequence represents a peptide compound of the invention having  
CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
CC invention are alpha-ketoamide peptide analogues. The peptides have  
CC virucide activity, and are useful for treating and in the manufacture of  
CC a medicament to treat disorders associated with HCV protease. A  
CC pharmaceutical composition comprising the peptide as an active ingredient  
CC is useful for treating disorders associated with hepatitis C virus.

XX SQ Sequence 11 AA;

Query Match 96.2%; Score 50; DB 23; Length 11;

Best Local Similarity 100.0%; Pred. No. 0.0011;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEVVPXGMSYS 11  
|||||  
Db 1 EEVVPXGMSYS 11

## RESULT 10

ABB80567

ID ABB80567 standard; peptide; 11 AA.

XX AC ABB80567;

XX DT 08-OCT-2002 (first entry)

DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #47.  
 XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
 KW virucide.  
 XX Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Modified-site 1 /note= "N-terminal acetyl"  
 FT Modified-site 6  
 FT Modified-site 6 /note= "(s,s)allothreonyl carbonyl residue forming a  
 FT keto-amide linkage with residue 7"  
 FT Modified-site 11  
 FT /note= "C-terminal amide"  
 FT  
 PN WO200208251-A2.  
 XX  
 XX 31-JAN-2002.  
 XX  
 XX 19-JUL-2001; 2001WO-US23169.  
 XX  
 XX 21-JUL-2000; 2000US-220101P.  
 XX (CORV-) CORVAS INT INC.  
 XX  
 XX Lim-wilby M, Levy OE, Brunck TK;  
 XX WPI; 2002-361643/39.  
 XX  
 XX Novel peptide compound having hepatitis C virus protease inhibitory  
 FT activity useful for treating disorders associated with hepatitis C  
 FT virus protease  
 XX  
 PS Claim 17; Page 65; 69pp; English.  
 XX  
 CC The sequence represents a peptide compound of the invention having  
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
 CC invention are alpha-ketoamide peptide analogues. The peptides have  
 CC virucide activity, and are useful for treating and in the manufacture of  
 CC a medicament to treat disorders associated with HCV protease. A  
 CC pharmaceutical composition comprising the peptide as an active ingredient  
 CC is useful for treating disorders associated with hepatitis C virus.  
 XX  
 SQ Sequence 11 AA;  
 XX  
 Query Match 96.2%; Score 50; DB 23; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 0.0011;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 EEVVPXGMSYS 11  
 Db 1 EEVVPXGMSYS 11  
 IIIIIIIIII  
 RESULT 11  
 ABB80568  
 ID ABB80568 standard; peptide; 11 AA.  
 AC ABB80568;  
 XX  
 XX 08-OCT-2002 (first entry)  
 DT  
 DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #48.  
 XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
 KW virucide.  
 XX Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Modified-site 1 /note= "N-terminal acetyl"  
 FT

FT Modified-site 6 /note= "Alpha-propionyl-glyciny-carbonyl residue forming  
 FT a keto-amide linkage with residue 7"  
 FT Modified-site 11  
 FT /note= "C-terminal amide"  
 XX  
 PN WO200208251-A2.  
 XX  
 XX 31-JAN-2002.  
 XX  
 XX 19-JUL-2001; 2001WO-US23169.  
 XX  
 XX 21-JUL-2000; 2000US-220101P.  
 XX (CORV-) CORVAS INT INC.  
 XX  
 XX Lim-wilby M, Levy OE, Brunck TK;  
 XX WPI; 2002-361643/39.  
 XX  
 XX Novel peptide compound having hepatitis C virus protease inhibitory  
 FT activity useful for treating disorders associated with hepatitis C  
 FT virus protease  
 XX  
 PS Claim 17; Page 65; 69pp; English.  
 XX  
 CC The sequence represents a peptide compound of the invention having  
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
 CC invention are alpha-ketoamide peptide analogues. The peptides have  
 CC virucide activity, and are useful for treating and in the manufacture of  
 CC a medicament to treat disorders associated with HCV protease. A  
 CC pharmaceutical composition comprising the peptide as an active ingredient  
 CC is useful for treating disorders associated with hepatitis C virus.  
 XX  
 SQ Sequence 11 AA;  
 XX  
 Query Match 96.2%; Score 50; DB 23; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 0.0011;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 EEVVPXGMSYS 11  
 Db 1 EEVVPXGMSYS 11  
 IIIIIIIIII  
 RESULT 12  
 ABB80524  
 ID ABB80524 standard; peptide; 11 AA.  
 XX  
 XX ABB80524;  
 XX  
 XX 08-OCT-2002 (first entry)  
 DT  
 DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #4.  
 XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
 KW virucide.  
 XX Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Modified-site 1 /note= "N-terminal acetyl"  
 FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with  
 FT residue 7"  
 FT Misc-difference 9 /note= "D-form residue"  
 FT Modified-site 11 /note= "C-terminal amide"  
 FT  
 PN WO200208251-A2.  
 XX

PD 31-JAN-2002.  
 XX  
 PF 19-JUL-2001; 2001WO-US23169.  
 XX  
 PR 21-JUL-2000; 2000US-220101P.  
 XX  
 PA (CORV-) CORVAS INT INC.  
 XX  
 PI Lim-wilby M, Levy OE, Brunck TK;  
 XX  
 DR WPI; 2002-361643/39.  
 XX  
 XX Novel peptide compound having hepatitis C virus protease inhibitory  
 PT activity useful for treating disorders associated with hepatitis C  
 PT virus protease -  
 XX  
 XX Claim 17; Page 64; 69pp; English.  
 PS  
 CC The sequence represents a peptide compound of the invention having  
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
 CC invention are alpha-ketoamide peptide analogues. The peptides have  
 CC virucide activity, and are useful for treating and in the manufacture of  
 CC a medicament to treat disorders associated with HCV protease. A  
 CC pharmaceutical composition comprising the peptide as an active ingredient  
 CC is useful for treating disorders associated with hepatitis C virus.  
 XX  
 XX Sequence 11 AA;  
 SQ  
 Query Match 88.5%; Score 46; DB 23; Length 11;  
 Best Local Similarity 90.9%; Pred. No. 0.0075;  
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 EEVVPXGMSYS 11  
 DB 1 EEVVPXGMDYS 11  
 RESULT 13  
 ABB80528  
 ID ABB80528 standard; peptide; 11 AA.  
 AC ABB80528;  
 XX  
 DT 08-OCT-2002 (first entry)  
 DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #8.  
 XX  
 KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
 KW virucide.  
 XX  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Modified-site 1 /note= "N-terminal acetyl"  
 FT Modified-site 6 /note= "D-form residue"  
 FT Modified-site 11 /note= "Norvalyl carbonyl forming keto-amide linkage with  
 FT residue 7"  
 FT Misc-difference 8 /note= "D-form residue"  
 FT Modified-site 9 /note= "D-form residue"  
 FT Modified-site 11 /note= "C-terminal amide"  
 XX  
 PN WO200208251-A2.  
 XX  
 PD 31-JAN-2002.  
 XX  
 PF 19-JUL-2001; 2001WO-US23169.  
 XX  
 PR 21-JUL-2000; 2000US-220101P.  
 XX  
 PA (CORV-) CORVAS INT INC.  
 XX  
 PI Lim-wilby M, Levy OE, Brunck TK;  
 XX  
 DR WPI; 2002-361643/39.  
 XX  
 PT Novel peptide compound having hepatitis C virus protease inhibitory  
 PT activity useful for treating disorders associated with hepatitis C

PI Lim-wilby M, Levy OE, Brunck TK;  
 XX  
 DR WPI; 2002-361643/39.  
 XX  
 PT Novel peptide compound having hepatitis C virus protease inhibitory  
 PT activity useful for treating disorders associated with hepatitis C  
 PT virus protease -  
 XX  
 PS Claim 17; Page 64; 69pp; English.  
 XX  
 CC The sequence represents a peptide compound of the invention having  
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
 CC invention are alpha-ketoamide peptide analogues. The peptides have  
 CC virucide activity, and are useful for treating and in the manufacture of  
 CC a medicament to treat disorders associated with HCV protease. A  
 CC pharmaceutical composition comprising the peptide as an active ingredient  
 CC is useful for treating disorders associated with hepatitis C virus.  
 XX  
 XX Sequence 11 AA;  
 SQ  
 Query Match 88.5%; Score 46; DB 23; Length 11;  
 Best Local Similarity 90.9%; Pred. No. 0.0075;  
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 EEVVPXGMSYS 11  
 DB 1 EEVVPXGMDYS 11  
 RESULT 14  
 ABB80529  
 ID ABB80529 standard; peptide; 11 AA.  
 AC ABB80529;  
 XX  
 DT 08-OCT-2002 (first entry)  
 DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #9.  
 XX  
 KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
 KW virucide.  
 XX  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Modified-site 1 /note= "N-terminal acetyl"  
 FT Modified-site 6 /note= "D-form residue"  
 FT Modified-site 11 /note= "Norvalyl carbonyl forming keto-amide linkage with  
 FT residue 7"  
 FT Misc-difference 8 /note= "D-form residue"  
 FT Misc-difference 9 /note= "D-form residue"  
 FT Modified-site 11 /note= "C-terminal amide"  
 XX  
 PN WO200208251-A2.  
 XX  
 PD 31-JAN-2002.  
 XX  
 PF 19-JUL-2001; 2001WO-US23169.  
 XX  
 PR 21-JUL-2000; 2000US-220101P.  
 XX  
 PA (CORV-) CORVAS INT INC.  
 XX  
 PI Lim-wilby M, Levy OE, Brunck TK;  
 XX  
 DR WPI; 2002-361643/39.  
 XX  
 PT Novel peptide compound having hepatitis C virus protease inhibitory  
 PT activity useful for treating disorders associated with hepatitis C

CC invention are alpha-ketoamide peptide analogues. The peptides have  
CC virucide activity, and are useful for treating and in the manufacture of  
CC a medicament to treat disorders associated with HCV protease. A  
CC pharmaceutical composition comprising the peptide as an active ingredient  
CC is useful for treating disorders associated with hepatitis C virus.  
XX  
XX Sequence 11 AA;  
SQ  
SQ

```

1 EEWXPXGMDYS 11
}|||||}
Db

```

RESULT 16  
ABB80562  
ID ABB80562 standard; peptide; 11 AA.  
XX  
AC ABB80562:

XX	
DT	
XX	
DE	08-OCT-2002 (first entry)
	Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #42.

XX	Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
KW	virucide.
KW	
XX	
XX	Synthetic.
OS	

OS	Synthetic.
XX	
XX	
FF	Key
FF	Location/Qualifiers
FF	Modified-site 1

[illegible]

FT	
FT	/note=Norvalyl carbonyl forming keto-amide linkage with residue 7 <sup>n</sup>
FT	Misc-difference 8

Misc-difference	8
/note= "D-form residue"	9

FT	/note= "D-form residue"
FT	8
FT	Modified-site
FT	8
FT	/note= "Oxymethionine"
FT	/note= "Oxymethionine"

FI	Moalled-site	8	/note=	"Oxymethionine"
FT				
FT	Misc-difference	9		
FT				

FT	note	oxymethionine
FT	Misc-difference 9	
FT	/note= "D-form residue"	

FT	/note= "D-form residue"
FT	11
Modified-site	
FT	
FT	/note= "C-terminal amide"
FT	
FT	/note= "C-terminal amide"

```

II
Modified-site
FT
FT
XX
./note= "C-terminal amide"

```

XX  
PN WO200208251-A2.

PN W0200208251-A2.  
XX  
PD 31 - JAN - 2002

XX  
31-JAN-2002.  
PD  
XX

19-JUL-2001; 2001WO-US23169.

19-JUL-2001; 2001WO-US23169.  
XX  
PF  
21-JUL-2000; 2000US-220101D  
PP

21-JUL-2000; 2000US-220101P.

41 JUL-2000; 200008-220101P.  
XX  
PA (CORV-) CORVAS INT INC.

PA (CORV-) CORVAS INT INC.  
XX  
XX

XX  
PI  
XX  
Lim-wilby M, Levy OE, Brunck TK;

PI LIM-WILDY M, LEVÝ OE, BRUNCK TK;  
XX  
DR WPI; 2002-361643/39.

DR 2002-361643/39.  
XX

XX  
PPT Novel peptide compound having hepatitis C virus protease inhibitory activity useful for treating disorders associated with hepatitis C

PT novel peptide compound having hepatitis C virus protease inhibitory activity useful for treating disorders associated with hepatitis C virus protease -

23. . . . .  
 24. . . . .  
 25. . . . .  
 26. . . . .  
 27. . . . .  
 28. . . . .  
 29. . . . .  
 30. . . . .  
 31. . . . .  
 32. . . . .  
 33. . . . .  
 34. . . . .  
 35. . . . .  
 36. . . . .  
 37. . . . .  
 38. . . . .  
 39. . . . .  
 40. . . . .  
 41. . . . .  
 42. . . . .  
 43. . . . .  
 44. . . . .  
 45. . . . .  
 46. . . . .  
 47. . . . .  
 48. . . . .  
 49. . . . .  
 50. . . . .  
 51. . . . .  
 52. . . . .  
 53. . . . .  
 54. . . . .  
 55. . . . .  
 56. . . . .  
 57. . . . .  
 58. . . . .  
 59. . . . .  
 60. . . . .  
 61. . . . .  
 62. . . . .  
 63. . . . .  
 64. . . . .  
 65. . . . .  
 66. . . . .  
 67. . . . .  
 68. . . . .  
 69. . . . .  
 70. . . . .  
 71. . . . .  
 72. . . . .  
 73. . . . .  
 74. . . . .  
 75. . . . .  
 76. . . . .  
 77. . . . .  
 78. . . . .  
 79. . . . .  
 80. . . . .  
 81. . . . .  
 82. . . . .  
 83. . . . .  
 84. . . . .  
 85. . . . .  
 86. . . . .  
 87. . . . .  
 88. . . . .  
 89. . . . .  
 90. . . . .  
 91. . . . .  
 92. . . . .  
 93. . . . .  
 94. . . . .  
 95. . . . .  
 96. . . . .  
 97. . . . .  
 98. . . . .  
 99. . . . .  
 100. . . . .

XX  
XX  
XX  
XX

Claim 17; Page 65; 69pp; English.

xx The sequence represents a peptide compound of the invention having  
CC Claim 1/; Page 65; 65pp; English.

The sequence represents a peptide compound of the invention having hepatitis C virus (HCV) protease inhibitory activity. The peptides of the

hepatitis C virus (HCV) protease inhibitory activity. The peptides of the invention are alpha-ketoamide peptide analogues. The peptides have antiviral activity and are useful for treating and preventing infection of

invention are alpha-ketoamide peptide analogues. The peptides have virucide activity, and are useful for treating and in the manufacture of a medicament to treat disorders associated with HCV protease. A

virulence activity, and are useful for treating and in the manufacture of a medicament to treat disorders associated with HCV protease. A pharmaceutical composition comprising the peptide as an active ingredient

pharmaceutical composition comprising the peptide as an active ingredient

CC is useful for treating disorders associated with hepatitis C virus.

SQ Sequence 11 AA;

Query Match 88.5%; Score 46; DB 23; Length 11;

Best Local Similarity 90.9%; Pred. No. 0.0075;

Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 EEVVPXGMSYS 11

DB 1 EEVVPXGMDYS 11

RESULT 17

ABB80523

ID ABB80523 standard; peptide; 11 AA.

XX AC ABB80523;

XX DT 08-OCT-2002 (first entry)

XX DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #3.

XX KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide; virucide.

XX OS Synthetic.

XX FH Key Location/Qualifiers

FT Modified-site 1

FT /note= "N-terminal acetyl"

FT Modified-site 6

FT /note= "Norvalyl carbonyl forming keto-amide linkage with residue 7"

FT Misc-difference 9

FT /note= "D-form residue"

FT Modified-site 11

FT /note= "C-terminal amide"

XX WO200208251-A2.

XX PD 31-JAN-2002.

XX PF 19-JUL-2001; 2001WO-US23169.

XX PR 21-JUL-2000; 2000US-220101P.

XX PA (CORV-) CORVAS INT INC.

XX PI Lim-wilby M, Levy OE, Brunck TK;

XX DR WPI; 2002-361643/39.

XX PT Novel peptide compound having hepatitis C virus protease inhibitory activity useful for treating disorders associated with hepatitis C virus protease

XX PS Claim 17; Page 64; 69pp; English.

XX CC The sequence represents a peptide compound of the invention having hepatitis C virus (HCV) protease inhibitory activity. The peptides of the invention are alpha-ketoamide peptide analogues. The peptides have virucide activity, and are useful for treating and in the manufacture of a medicament to treat disorders associated with HCV protease. A pharmaceutical composition comprising the peptide as an active ingredient is useful for treating disorders associated with hepatitis C virus.

XX SQ Sequence 11 AA;

Query Match

Best Local Similarity 86.5%; Score 45; DB 23; Length 11;

Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 EEVVPXGMSYS 11

DB 1 EEVVPXGMHYS 11

RESULT 18

ABB80527

ID ABB80527 standard; peptide; 11 AA.

XX AC ABB80527;

XX DT 08-OCT-2002 (first entry)

XX DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #7.

XX KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide; virucide.

XX OS Synthetic.

XX FH Key Location/Qualifiers

FT Modified-site 1

FT /note= "N-terminal acetyl"

FT Modified-site 6

FT /note= "Norvalyl carbonyl forming keto-amide linkage with residue 7"

FT Misc-difference 8

FT /note= "D-form residue"

FT Modified-site 11

FT /note= "C-terminal amide"

XX WO200208251-A2.

XX PD 31-JAN-2002.

XX PF 19-JUL-2001; 2001WO-US23169.

XX PR 21-JUL-2000; 2000US-220101P.

XX PA (CORV-) CORVAS INT INC.

XX PI Lim-wilby M, Levy OE, Brunck TK;

XX DR WPI; 2002-361643/39.

XX PT Novel peptide compound having hepatitis C virus protease inhibitory activity useful for treating disorders associated with hepatitis C virus protease

XX PS Claim 17; Page 64; 69pp; English.

XX CC The sequence represents a peptide compound of the invention having hepatitis C virus (HCV) protease inhibitory activity. The peptides of the invention are alpha-ketoamide peptide analogues. The peptides have virucide activity, and are useful for treating and in the manufacture of a medicament to treat disorders associated with HCV protease. A pharmaceutical composition comprising the peptide as an active ingredient is useful for treating disorders associated with hepatitis C virus.

XX SQ Sequence 11 AA;

Query Match

Best Local Similarity 86.5%; Score 45; DB 23; Length 11;

Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 EEVVPXGMSYS 11

DB 1 EEVVPXGMHYS 11

RESULT 19

ABB80535

ID ABB80535 standard; peptide; 11 AA.





FT Misc-difference 8  
 FT FT /note= "D-form residue"  
 FT Modified-site 11  
 FT FT /note= "C-terminal amide"

PN WO200208251-A2.

XX 31-JAN-2002.

XX 19-JUL-2001; 2001WO-US23169.

XX 21-JUL-2000; 2000US-220101P.

XX (CORV-) CORVAS INT INC.

XX Lim-wilby M, Levy OE, Brunck TK;

XX WPI; 2002-361643/39.

XX Novel peptide compound having hepatitis C virus protease inhibitory  
 PT activity useful for treating disorders associated with hepatitis C  
 PT virus protease

XX Claim 17; Page 65; 69pp; English.

XX The sequence represents a peptide compound of the invention having  
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
 CC invention are alpha-ketoamide peptide analogues. The peptides have  
 CC virucide activity, and are useful for treating and in the manufacture of  
 CC a medicament to treat disorders associated with HCV protease. A  
 CC pharmaceutical composition comprising the peptide as an active ingredient  
 CC is useful for treating disorders associated with hepatitis C virus.

XX Sequence 11 AA;

Query Match 86.5%; Score 45; DB 23; Length 11;

Best Local Similarity 90.9%; Pred. No. 0.012;

Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXGMSYS 11

DB 1 EEVVPXGQSYS 11

RESULT 22

ABB80540

ID ABB80540 standard; peptide; 11 AA.

XX ABB80540;

XX 08-OCT-2002 (first entry)

XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #20.

XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
 KW virucide.

XX Synthetic.

XX Key Location/Qualifiers

FT Modified-site 1

FT /note= "N-terminal acetyl"

FT Modified-site 6

FT /note= "Norvalyl carbonyl forming keto-amide linkage with  
 FT residue 7"

FT Misc-difference 8

FT /note= "D-form residue"

FT Misc-difference 9

FT /note= "D-form residue"

FT Modified-site 11

FT /note= "C-terminal amide"

XX WO200208251-A2.

XX

PD 31-JAN-2002.

XX 19-JUL-2001; 2001WO-US23169.

XX 21-JUL-2000; 2000US-220101P.

XX (CORV-) CORVAS INT INC.

XX Lim-wilby M, Levy OE, Brunck TK;

XX WPI; 2002-361643/39.

XX Novel peptide compound having hepatitis C virus protease inhibitory  
 PT activity useful for treating disorders associated with hepatitis C  
 PT virus protease

XX Claim 17; Page 65; 69pp; English.

XX The sequence represents a peptide compound of the invention having  
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
 CC invention are alpha-ketoamide peptide analogues. The peptides have  
 CC virucide activity, and are useful for treating and in the manufacture of  
 CC a medicament to treat disorders associated with HCV protease. A  
 CC pharmaceutical composition comprising the peptide as an active ingredient  
 CC is useful for treating disorders associated with hepatitis C virus.

XX Sequence 11 AA;

Query Match 86.5%; Score 45; DB 23; Length 11;

Best Local Similarity 90.9%; Pred. No. 0.012;

Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXGMSYS 11

DB 1 EEVVPXGQSYS 11

RESULT 23

ABB80558

ID ABB80558 standard; peptide; 11 AA.

XX ABB80558;

XX 08-OCT-2002 (first entry)

XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #38.

XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
 KW virucide.

XX Synthetic.

XX Key Location/Qualifiers

FT Modified-site 1

FT /note= "N-terminal acetyl"

FT Modified-site 6

FT /note= "Norvalyl carbonyl forming keto-amide linkage with  
 FT residue 7"

FT Modified-site 8

FT /note= "Oxymethionine"

FT Modified-site 11

FT /note= "C-terminal amide"

XX WO200208251-A2.

XX 31-JAN-2002.

XX 19-JUL-2001; 2001WO-US23169.

XX 21-JUL-2000; 2000US-220101P.

XX (CORV-) CORVAS INT INC.

XX PI Lim-wilby M, Levy OE, Brunck TK;  
 XX WPI; 2002-361643/39.  
 XX Novel peptide compound having hepatitis C virus protease inhibitory  
 PT activity useful for treating disorders associated with hepatitis C  
 PT virus protease -  
 XX  
 XX Claim 17; Page 65; 69pp; English.  
 CC The sequence represents a peptide compound of the invention having  
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
 CC invention are alpha-ketoamide peptide analogues. The peptides have  
 CC virucide activity, and are useful for treating and in the manufacture of  
 CC a medicament to treat disorders associated with HCV protease. A  
 CC pharmaceutical composition comprising the peptide as an active ingredient  
 CC is useful for treating disorders associated with hepatitis C virus.  
 XX SQ Sequence 11 AA;  
 Query Match 86.5%; Score 45; DB 23; Length 11;  
 Best Local Similarity 90.9%; Pred. No. 0.012;  
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 OY 1 EEVVPXGMSYS 11  
 Db 1 EEVVPXGMHYS 11  
 RESULT 24  
 ABB80560  
 ID ABB80560 standard; peptide; 11 AA.  
 AC ABB80560;  
 XX 08-OCT-2002 (first entry)  
 XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #40.  
 DE Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
 KW virucide.  
 KW Synthetic.  
 OS Key Location/Qualifiers  
 XX Modified-site 1 /note= "N-terminal acetyl"  
 FT Modified-site 6  
 FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with  
 FT residue 7"  
 FT Misc-difference 8 /note= "D-form residue"  
 FT Modified-site 8 /note= "Oxymethionine"  
 FT Misc-difference 9 /note= "D-form residue"  
 FT Modified-site 11 /note= "C-terminal amide"  
 FT WO200208251-A2.  
 XX 31-JAN-2002.  
 XX 19-JUL-2001; 2001WO-US23169.  
 XX 21-JUL-2000; 2000US-220101P.  
 XX (CORV-) CORVAS INT INC.  
 XX Lim-wilby M, Levy OE, Brunck TK;  
 XX WPI; 2002-361643/39.  
 XX Novel peptide compound having hepatitis C virus protease inhibitory  
 PT activity useful for treating disorders associated with hepatitis C  
 PT virus protease -  
 XX  
 XX Claim 17; Page 65; 69pp; English.  
 CC The sequence represents a peptide compound of the invention having  
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
 CC invention are alpha-ketoamide peptide analogues. The peptides have  
 CC virucide activity, and are useful for treating and in the manufacture of  
 CC a medicament to treat disorders associated with HCV protease. A  
 CC pharmaceutical composition comprising the peptide as an active ingredient  
 CC is useful for treating disorders associated with hepatitis C virus.  
 XX SQ Sequence 11 AA;  
 Query Match 86.5%; Score 45; DB 23; Length 11;  
 Best Local Similarity 90.9%; Pred. No. 0.012;  
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 OY 1 EEVVPXGMSYS 11  
 Db 1 EEVVPXGMHYS 11

XX Novel peptide compound having hepatitis C virus protease inhibitory  
 PT activity useful for treating disorders associated with hepatitis C  
 PT virus protease -  
 XX  
 XX Claim 17; Page 65; 69pp; English.  
 CC The sequence represents a peptide compound of the invention having  
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
 CC invention are alpha-ketoamide peptide analogues. The peptides have  
 CC virucide activity, and are useful for treating and in the manufacture of  
 CC a medicament to treat disorders associated with HCV protease. A  
 CC pharmaceutical composition comprising the peptide as an active ingredient  
 CC is useful for treating disorders associated with hepatitis C virus.  
 XX SQ Sequence 11 AA;  
 Query Match 86.5%; Score 45; DB 23; Length 11;  
 Best Local Similarity 90.9%; Pred. No. 0.012;  
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 OY 1 EEVVPXGMSYS 11  
 Db 1 EEVVPXGMHYS 11  
 RESULT 25  
 ABB80544  
 ID ABB80544 standard; peptide; 11 AA.  
 AC ABB80544;  
 XX 08-OCT-2002 (first entry)  
 XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #24.  
 DE Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
 KW virucide.  
 KW Synthetic.  
 OS Key Location/Qualifiers  
 XX Modified-site 1 /note= "N-terminal acetyl"  
 FT Modified-site 6  
 FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with  
 FT residue 7"  
 FT Modified-site 11 /note= "C-terminal amide"  
 FT WO200208251-A2.  
 XX 31-JAN-2002.  
 XX 19-JUL-2001; 2001WO-US23169.  
 XX 21-JUL-2000; 2000US-220101P.  
 XX (CORV-) CORVAS INT INC.  
 XX Lim-wilby M, Levy OE, Brunck TK;  
 XX WPI; 2002-361643/39.  
 XX Novel peptide compound having hepatitis C virus protease inhibitory  
 PT activity useful for treating disorders associated with hepatitis C  
 PT virus protease -  
 XX  
 XX Claim 17; Page 65; 69pp; English.  
 CC The sequence represents a peptide compound of the invention having  
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
 CC invention are alpha-ketoamide peptide analogues. The peptides have

CC virucide activity, and are useful for treating and in the manufacture of  
CC a medicament to treat disorders associated with HCV protease. A  
CC pharmaceutical composition comprising the peptide as an active ingredient  
CC is useful for treating disorders associated with hepatitis C virus.

XX  
SQ Sequence 11 AA;

Query Match 84.6%; Score 44; DB 23; Length 11;  
Best Local Similarity 90.9%; Pred. No. 0.019;  
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 EEVVPXGMSYS 11  
|||||||  
Db 1 EEVVPXGTSYS 11

Search completed: June 10, 2003, 13:39:06  
Job time : 32.3571 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2003, 13:31:45 ; Search time 9.64286 Seconds  
(without alignments)  
33.564 Million cell updates/sec

Title: US-09-909-164-6  
Perfect score: 52  
Sequence: 1 BEVWPXGMSYS 11

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 75 summaries

Database : Issued Patents.AA.\*

1: /cgn2.6/ptodata/1/1aa/5A\_COMB.pep.\*  
2: /cgn2.6/ptodata/1/1aa/5B\_COMB.pep.\*  
3: /cgn2.6/ptodata/1/1aa/6A\_COMB.pep.\*  
4: /cgn2.6/ptodata/1/1aa/6B\_COMB.pep.\*  
5: /cgn2.6/ptodata/1/1aa/PTUS\_COMB.pep.\*  
6: /cgn2.6/ptodata/1/1aa/backfiles1.pep.\*

pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	34	65.4	947	4	US-09-228-986-73
2	33	63.5	45	2	US-08-637-759B-236
3	33	63.5	45	3	US-08-871-355A-236
4	33	63.5	45	4	US-09-201-945-236
5	33	63.5	65	6	5177197-51
6	33	63.5	410	6	5177197-1
7	33	63.5	1394	6	5177197-30
8	32	61.5	10	4	US-09-357-952-66
9	32	61.5	10	4	US-09-521-650-66
10	32	61.5	10	4	US-09-168-888-66
11	32	61.5	102	2	US-08-580-988A-23
12	32	61.5	152	2	US-08-460-694-4
13	32	61.5	152	3	US-08-460-744-4
14	32	61.5	152	3	US-07-667-711B-4
15	32	61.5	173	1	US-08-193-977-7
16	32	61.5	189	2	US-08-464-517-21
17	32	61.5	189	2	US-08-246-361A-21
18	32	61.5	189	3	US-08-463-772-21
19	32	61.5	189	5	PCT-US93-05000-21
20	32	61.5	236	2	US-08-464-517-22
21	32	61.5	236	2	US-08-246-361A-22
22	32	61.5	236	3	US-08-463-772-22
23	32	61.5	236	5	PCT-US93-05000-22
24	32	61.5	280	2	US-08-464-517-6
25	32	61.5	280	3	US-08-463-772-6
26	32	61.5	289	2	US-08-246-361A-4
27	32	61.5	289	5	PCT-US93-05000-4

28	32	61.5	291	5	PCT-US93-05000-6	Sequence 6, Appl
29	32	61.5	292	2	US-08-464-517-23	Sequence 23, Appl
30	32	61.5	292	2	US-08-246-361A-6	Sequence 6, Appl
31	32	61.5	292	2	US-08-246-361A-23	Sequence 23, Appl
32	32	61.5	292	3	US-08-463-772-23	Sequence 23, Appl
33	32	61.5	292	5	PCT-US93-05000-23	Sequence 23, Appl
34	32	61.5	295	1	US-07-947-120-8	Sequence 8, Appl
35	32	61.5	295	1	US-08-472-893A-8	Sequence 8, Appl
36	32	61.5	295	2	US-08-460-694-2	Sequence 2, Appl
37	32	61.5	295	2	US-08-464-517-19	Sequence 19, Appl
38	32	61.5	295	2	US-08-464-517-20	Sequence 20, Appl
39	32	61.5	295	2	US-08-246-361A-19	Sequence 19, Appl
40	32	61.5	295	2	US-08-246-361A-20	Sequence 20, Appl
41	32	61.5	295	3	US-08-463-772-19	Sequence 19, Appl
42	32	61.5	295	3	US-08-463-772-20	Sequence 20, Appl
43	32	61.5	295	3	US-08-460-744-2	Sequence 2, Appl
44	32	61.5	295	3	US-07-667-711B-2	Sequence 2, Appl
45	32	61.5	295	3	US-08-947-492-8	Sequence 8, Appl
46	32	61.5	295	5	PCT-US93-05000-2	Sequence 2, Appl
47	32	61.5	295	5	PCT-US93-05000-19	Sequence 19, Appl
48	32	61.5	295	5	PCT-US93-05000-20	Sequence 20, Appl
49	32	61.5	309	2	US-08-464-517-4	Sequence 4, Appl
50	32	61.5	309	3	US-08-463-772-4	Sequence 4, Appl
51	32	61.5	529	4	US-09-240-639-4	Sequence 4, Appl
52	32	61.5	618	2	US-08-770-761A-3	Sequence 3, Appl
53	32	61.5	647	2	US-08-770-761A-8	Sequence 8, Appl
54	32	61.5	660	2	US-08-770-761A-2	Sequence 2, Appl
55	32	61.5	662	2	US-08-770-761A-5	Sequence 5, Appl
56	32	61.5	705	2	US-08-770-761A-7	Sequence 7, Appl
57	32	61.5	819	2	US-08-464-517-7	Sequence 7, Appl
58	32	61.5	819	2	US-08-246-361A-7	Sequence 7, Appl
59	32	61.5	819	3	US-08-463-772-7	Sequence 7, Appl
60	32	61.5	819	5	PCT-US93-05000-7	Sequence 7, Appl
61	31	59.6	59	4	US-08-963-851-14	Sequence 14, Appl
62	31	59.6	622	2	US-08-459-146-2	Sequence 2, Appl
63	31	59.6	622	2	US-08-459-065-2	Sequence 2, Appl
64	30	57.7	13	4	US-09-288-391-22	Sequence 22, Appl
65	30	57.7	13	4	US-09-288-391-23	Sequence 23, Appl
66	30	57.7	117	4	US-08-444-818-44	Sequence 44, Appl
67	30	57.7	121	4	US-09-152-060-68	Sequence 68, Appl
68	30	57.7	121	4	US-09-152-060-85	Sequence 85, Appl
69	30	57.7	122	3	US-08-879-995A-1	Sequence 1, Appl
70	30	57.7	122	3	US-09-215-096-1	Sequence 1, Appl
71	30	57.7	132	4	US-08-444-818-52	Sequence 52, Appl
72	30	57.7	159	2	US-08-844-086-4	Sequence 4, Appl
73	30	57.7	159	3	US-09-018-211-4	Sequence 4, Appl
74	30	57.7	211	5	PCT-US94-04174-18	Sequence 18, Appl
75	30	57.7	241	3	US-08-834-776A-2	Sequence 2, Appl

#### ALIGNMENTS

RESULT 1  
US-09-228-986-73  
; Sequence 73, Application US/09228986  
; Patent No. 6359198  
; GENERAL INFORMATION:  
; APPLICANT: Strabala, Timothy  
; APPLICANT: Nieuwenhuizen, Niels  
; TITLE OF INVENTION: Compositions Isolated from Plant Cells  
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signalling  
; FILE REFERENCE: 11000/1020  
; CURRENT APPLICATION NUMBER: US/09/228,986  
; CURRENT FILING DATE: 1999-01-12  
; NUMBER OF SEQ ID NOS: 130  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 73  
; LENGTH: 947  
; TYPE: PRT  
; ORGANISM: Pinus radiata  
US-09-228-986-73

Query Match 65.4%; Score 34; DB 4; Length 947;  
 Best Local Similarity 66.7%; Pred. No. 1e-02;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMSYS 11  
 I: I I I I I  
 Db 686 VMPGSIYS 694

## RESULT 2

US-08-637-759B-236  
 ; Sequence 236, Application US/08637759B  
 ; Patent No. 5876931

## GENERAL INFORMATION:

APPLICANT: David William Holden  
 TITLE OF INVENTION: Identification of Genes  
 NUMBER OF SEQUENCES: 501  
 CORRESPONDENCE ADDRESS:

ADDRESSEE: Patrea L. Pabst  
 STREET: 2800 One Atlantic Center  
 CITY: Atlanta  
 STATE: Georgia  
 COUNTRY: USA  
 ZIP: 30309-3450

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/637,759B  
 FILING DATE: 03-MAY-1996  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: PCT/GB95/02875  
 FILING DATE: 11-DEC-1995

## ATTORNEY/AGENT INFORMATION:

NAME: Pabst, Patrea L.  
 REGISTRATION NUMBER: 31,284  
 REFERENCE/DOCKET NUMBER: RPMS 101  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (404) 873-8794  
 TELEFAX: (404) 873-8795

## INFORMATION FOR SEQ ID NO: 236:

SEQUENCE CHARACTERISTICS:  
 LENGTH: 45 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 HYPOTHETICAL: NO

US-08-637-759B-236

Query Match 63.5%; Score 33; DB 2; Length 45;  
 Best Local Similarity 60.0%; Pred. No. 5.1;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10  
 I: I I I I I  
 Db 1 EEISPLGWSY 10

## RESULT 3

US-08-871-355A-236  
 ; Sequence 236, Application US/08871355A  
 ; Patent No. 6015669

## GENERAL INFORMATION:

APPLICANT: David William Holden  
 TITLE OF INVENTION: Identification of Genes  
 NUMBER OF SEQUENCES: 501  
 CORRESPONDENCE ADDRESS:

ADDRESSEE: Patrea L. Pabst  
 STREET: 2800 One Atlantic Center  
 CITY: Atlanta  
 STATE: Georgia  
 COUNTRY: USA  
 ZIP: 30309-3450  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/871,355A  
 FILING DATE: 09-JUN-1997  
 CLASSIFICATION: 435

## PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/GB95/02875  
 FILING DATE: 11-DEC-1995  
 CLASSIFICATION: 435

## ATTORNEY/AGENT INFORMATION:

NAME: Pabst, Patrea L.  
 REGISTRATION NUMBER: 31,284  
 REFERENCE/DOCKET NUMBER: RPMS 101 CON  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (404) 873-8794  
 TELEFAX: (404) 873-8795

## INFORMATION FOR SEQ ID NO: 236:

SEQUENCE CHARACTERISTICS:  
 LENGTH: 45 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 HYPOTHETICAL: NO

US-08-871-355A-236

Query Match 63.5%; Score 33; DB 3; Length 45;  
 Best Local Similarity 60.0%; Pred. No. 5.1;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10  
 I: I I I I I  
 Db 1 EEISPLGWSY 10

## RESULT 4

US-09-201-945-236  
 ; Sequence 236, Application US/09201945  
 ; Patent No. 6342215

## GENERAL INFORMATION:

APPLICANT: David William Holden  
 TITLE OF INVENTION: Identification of Genes  
 NUMBER OF SEQUENCES: 501  
 CORRESPONDENCE ADDRESS:

ADDRESSEE: Patrea L. Pabst  
 STREET: 2800 One Atlantic Center  
 CITY: Atlanta  
 STATE: Georgia  
 COUNTRY: USA  
 ZIP: 30309-3450

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/201,945  
 FILING DATE:  
 CLASSIFICATION:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/637,759

; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Pabst, Patrea L.  
; REGISTRATION NUMBER: 31,284  
; REFERENCE/DOCKET NUMBER: RPS 101  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (404) 873-8794  
; TELEFAX: (404) 873-8795  
; INFORMATION FOR SEQ ID NO: 236:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 45 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHEetical: NO  
; US-09-201-945-236

Query Match 63.5%; Score 33; DB 4; Length 45;  
Best Local Similarity 60.0%; Pred. No. 5.1;  
Matches 6; Conservative 1; Mismatches 3; Indels 3; Gaps 0;

QY 1 EEVVPXGMSY 10  
; :|: | | | |  
Db 1 EEISPLGWSY 10

## RESULT 5

5177197-51  
; Patent No. 5177197  
; APPLICANT: KANZAKI, TETSUTO; OLOFSSON, ANDERS; MOREN, ANITA;  
; WERNSTEDT, CHRISTER; HELLMAN, ULF; MIYAZONO, KOHEI; CLAESSON-WELSH,  
; LENA; HELDIN, CARL-HENRIK  
; TITLE OF INVENTION: ISOLATED NUCLEOTIDE SEQUENCE EXPRESSING  
; HUMAN TRANSFORMING GROWTH FACTOR-BETAL-BINDING PROTEIN  
; NUMBER OF SEQUENCES: 53  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/487,343  
; FILING DATE: 27-FEB-1990  
; SEQ ID NO: 51:  
; LENGTH: 65  
5177197-51

Query Match 63.5%; Score 33; DB 6; Length 65;  
Best Local Similarity 45.5%; Pred. No. 7.8;  
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSYS 11  
; :|: | | | |  
Db 52 KEICPGMGYT 62

## RESULT 6

5177197-1  
; Patent No. 5177197  
; APPLICANT: KANZAKI, TETSUTO; OLOFSSON, ANDERS; MOREN, ANITA;  
; WERNSTEDT, CHRISTER; HELLMAN, ULF; MIYAZONO, KOHEI; CLAESSON-WELSH,  
; LENA; HELDIN, CARL-HENRIK  
; TITLE OF INVENTION: ISOLATED NUCLEOTIDE SEQUENCE EXPRESSING  
; HUMAN TRANSFORMING GROWTH FACTOR-BETAL-BINDING PROTEIN  
; NUMBER OF SEQUENCES: 53  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/487,343  
; FILING DATE: 27-FEB-1990  
; SEQ ID NO: 1:  
; LENGTH: 410  
5177197-1

Query Match 63.5%; Score 33; DB 6; Length 410;  
Best Local Similarity 45.5%; Pred. No. 63;  
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSYS 11  
; :|: | | | |  
Db 399 KEICPGMGYT 409

## RESULT 7

5177197-30  
; Patent No. 5177197  
; APPLICANT: KANZAKI, TETSUTO; OLOFSSON, ANDERS; MOREN, ANITA;  
; WERNSTEDT, CHRISTER; HELLMAN, ULF; MIYAZONO, KOHEI; CLAESSON-WELSH,  
; LENA; HELDIN, CARL-HENRIK  
; TITLE OF INVENTION: ISOLATED NUCLEOTIDE SEQUENCE EXPRESSING  
; HUMAN TRANSFORMING GROWTH FACTOR-BETAL-BINDING PROTEIN  
; NUMBER OF SEQUENCES: 53  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/487,343  
; FILING DATE: 27-FEB-1990  
; SEQ ID NO: 30:  
; LENGTH: 1394  
5177197-30

Query Match 63.5%; Score 33; DB 6; Length 1394;  
Best Local Similarity 45.5%; Pred. No. 2.6e+02;  
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSYS 11  
; :|: | | | |  
Db 399 KEICPGMGYT 409

## RESULT 8

US-09-357-952-66  
; Sequence 66, Application US/09357952  
; Patent No. 6248904  
; GENERAL INFORMATION:  
; APPLICANT: Zhang, Han-Zhong  
; APPLICANT: Cai, Sui Xiong  
; APPLICANT: Drewe, John A.  
; APPLICANT: Yang, Wu  
; TITLE OF INVENTION: No. 6248904el Fluorescence Dyes and Their Applications for W  
; TITLE OF INVENTION: Fluorescence Screening Assays for Caspases, Peptidases, Prot  
; FILE REFERENCE: 1735.0030001  
; CURRENT APPLICATION NUMBER: US/09/357,952  
; CURRENT FILING DATE: 1999-07-21  
; EARLIER APPLICATION NUMBER: US 60/093,642  
; EARLIER FILING DATE: 21-JUL-1998  
; NUMBER OF SEQ ID NOS: 139  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 66  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
; OTHER INFORMATION: Peptide  
US-09-357-952-66

Query Match 61.5%; Score 32; DB 4; Length 10;  
Best Local Similarity 50.0%; Pred. No. 1.5;  
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10  
; :|: | | | |  
Db 1 DDIVPCMSY 10

## RESULT 9

US-09-521-650-66  
; Sequence 66, Application US/09521650  
; Patent No. 6335429  
; GENERAL INFORMATION:  
; APPLICANT: Weber, Eckard

APPLICANT: Cai, Sui Xiong  
APPLICANT: Keana, John F.W.  
APPLICANT: Drewe, John A.  
APPLICANT: Zhang, Han-Zhong  
TITLE OF INVENTION: No. 6335429el Fluorogenic or Fluorescent Reporter Molecules and  
TITLE OF INVENTION: Their Applications for Whole-Cell Fluorescence  
TITLE OF INVENTION: Screening Assays for Caspases and Other Enzymes and the  
TITLE OF INVENTION: Use Thereof  
FILE REFERENCE: 1735.0290002  
CURRENT APPLICATION NUMBER: US/09/521,650  
CURRENT FILING DATE: 2000-03-08  
EARLIER APPLICATION NUMBER: 09/168,888  
EARLIER FILING DATE: 1998-10-09  
EARLIER APPLICATION NUMBER: US 60/061,582  
EARLIER FILING DATE: 1997-10-10  
EARLIER APPLICATION NUMBER: US 09/033,661  
EARLIER FILING DATE: 1998-03-03  
NUMBER OF SEQ ID NOS: 142  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 66  
LENGTH: 10  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
OTHER INFORMATION: Peptide  
US-09-521-650-66

Query Match 61.5%; Score 32; DB 4; Length 10;  
Best Local Similarity 50.0%; Pred. No. 1.5;  
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10  
::||| |||  
Db 1 DDIVPCMSY 10

RESULT 10  
US-09-168-888-66  
Sequence 66, Application US/09168888  
Patent No. 6342611  
GENERAL INFORMATION:  
APPLICANT: Weber, Eckard  
APPLICANT: Cai, Sui Xiong  
APPLICANT: Keana, John F.W.  
APPLICANT: Drewe, John A.  
APPLICANT: Zhang, Han-Zhong  
TITLE OF INVENTION: No. 6342611el Fluorogenic or Fluorescent Reporter Molecules and  
TITLE OF INVENTION: Their Applications for Whole-Cell Fluorescence  
TITLE OF INVENTION: Screening Assays for Caspases and Other Enzymes and the  
TITLE OF INVENTION: Use Thereof  
FILE REFERENCE: 1735.0290002  
CURRENT APPLICATION NUMBER: US/09/168,888  
CURRENT FILING DATE: 1998-10-09  
EARLIER APPLICATION NUMBER: US 60/061,582  
EARLIER FILING DATE: 1997-10-10  
EARLIER APPLICATION NUMBER: US 09/033,661  
EARLIER FILING DATE: 1998-03-03  
NUMBER OF SEQ ID NOS: 142  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 66  
LENGTH: 10  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
OTHER INFORMATION: Peptide  
US-09-168-888-66

Query Match 61.5%; Score 32; DB 4; Length 10;  
Best Local Similarity 50.0%; Pred. No. 1.5;  
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10  
::||| |||  
Db 1 DDIVPCMSY 10

RESULT 11  
US-08-580-988A-23  
Sequence 23, Application US/08580988A  
Patent No. 5856161  
GENERAL INFORMATION:  
APPLICANT: Aggarwal et al.  
TITLE OF INVENTION: Tumor Necrosis Factor  
TITLE OF INVENTION: Receptor-I-Associated Protein Kinase And Methods  
TITLE OF INVENTION: For Its Use  
NUMBER OF SEQUENCES: 27  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dr. Benjamin A. Adler  
STREET: 8011 Candle Lane  
CITY: Houston  
STATE: Texas  
COUNTRY: USA  
ZIP: 77071  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 1.44 MD floppy disk  
COMPUTER: Apple Macintosh  
OPERATING SYSTEM: Macintosh  
SOFTWARE: Microsoft Word for Macintosh  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/580,988A  
FILING DATE: January 3, 1996  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Benjamin Aaron Adler, Ph.D., J.D.  
REGISTRATION NUMBER: 35,423  
REFERENCE/DOCKET NUMBER: D5721CIP2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 713-777-2321  
TELEFAX: 713-777-6908  
INFORMATION FOR SEQ ID NO: 23:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 102 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE:  
DESCRIPTION: protein  
HYPOTHETICAL: no  
ANTI-SENSE: no  
FRAGMENT TYPE: internal  
ORIGINAL SOURCE:  
US-08-580-988A-23

Query Match 61.5%; Score 32; DB 2; Length 102;  
Best Local Similarity 60.0%; Pred. No. 21;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10  
||| |||  
Db 24 EEVPLAMNY 33

RESULT 12  
US-08-460-694-4  
Sequence 4, Application US/08460694  
Patent No. 5858655  
GENERAL INFORMATION:  
APPLICANT: Arnold, Andrew  
TITLE OF INVENTION: PRAD1 Cyclin and its cDNA  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:



ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
STREET: 1100 New York Avenue, N.W., Suite 600  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/460,694  
FILING DATE: 02-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: McConathy, Evelyn H.  
REGISTRATION NUMBER: 35,279  
REFERENCE/DOCKET NUMBER: 0609.4070002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-2600  
TELEFAX: 202-371-2540  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 152 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-460-694-4

Query Match 61.5%; Score 32; DB 2; Length 152;  
Best Local Similarity 60.0%; Pred. No. 33;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 EEVVPXGMSY 10  
||| | | |  
Db 20 EEVFPPLAMNY 29

RESULT 13  
US-08-460-744-4  
Sequence 4, Application US/08460744  
Patent No. 6107541  
GENERAL INFORMATION:  
APPLICANT: Arnold, Andrew  
TITLE OF INVENTION: Pradi Cyclin and its cDNA  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
STREET: 1100 New York Avenue, N.W., Suite 600  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/460,744  
FILING DATE: 02-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: McConathy, Evelyn H.  
REGISTRATION NUMBER: 35,279  
REFERENCE/DOCKET NUMBER: 0609.4070005  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-2600  
TELEFAX: 202-371-2540  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:

LENGTH: 152 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-460-744-4

Query Match 61.5%; Score 32; DB 3; Length 152;  
Best Local Similarity 60.0%; Pred. No. 33;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 EEVVPXGMSY 10  
||| | | |  
Db 20 EEVFPPLAMNY 29

RESULT 14  
US-07-667-711B-4  
Sequence 4, Application US/07667711B  
Patent No. 6110700  
GENERAL INFORMATION:  
APPLICANT: ARNOLD, ANDREW  
TITLE OF INVENTION: Pradi Cyclin and its cDNA  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
STREET: 1100 NEW YORK AVE., NW, SUITE 600  
CITY: WASHINGTON  
STATE: DC  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/667,711B  
FILING DATE: 11-MAR-1991  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: MCPHAIL, DONALD R.  
REGISTRATION NUMBER: 35,811  
REFERENCE/DOCKET NUMBER: 0609.4070000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 371-2600  
TELEFAX: (202) 371-2540  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 152 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: not relevant  
MOLECULE TYPE: peptide  
US-07-667-711B-4

Query Match 61.5%; Score 32; DB 3; Length 152;  
Best Local Similarity 60.0%; Pred. No. 33;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 EEVVPXGMSY 10  
||| | | |  
Db 20 EEVFPPLAMNY 29

RESULT 15  
US-08-193-977-7  
Sequence 7, Application US/08193977  
Patent No. 5625031  
GENERAL INFORMATION:  
APPLICANT: WEBSTER, KEVIN R.  
APPLICANT: COLEMAN, KEVIN G.  
TITLE OF INVENTION: PEPTIDE INHIBITORS OF THE P33CDK2 AND

;; TITLE OF INVENTION: P34CDC2 CELL CYCLE REGULATORY KINASES AND HUMAN  
;; TITLE OF INVENTION: PAPILLOMAVIRUS E7 ONCOPROTEIN  
;; NUMBER OF SEQUENCES: 34  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: REED & ROBINS  
;; STREET: 635 BRYANT STREET  
;; CITY: PALO ALTO  
;; STATE: CALIFORNIA  
;; COUNTRY: UNITED STATES OF AMERICA  
;; ZIP: 94301  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/193,977  
;; FILING DATE: 08-FEB-1994  
;; CLASSIFICATION: 530  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: ROBINS, ROBERTA L.  
;; REGISTRATION NUMBER: 33,208  
;; REFERENCE/DOCKET NUMBER: 5998-0016  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (415) 617-8959  
;; TELEFAX: (415) 327-3231  
;; INFORMATION FOR SEQ ID NO: 7:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 173 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
;; US-08-193-977-7

Query Match 61.5%; Score 32; DB 1; Length 173;  
Best Local Similarity 60.0%; Pred. No. 38;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10  
||| | | |  
DB 55 EEVFPPLAMNY 64

RESULT 16  
US-08-464-517-21  
;; Sequence 21, Application US/08464517  
;; Patent No. 5869640  
;; GENERAL INFORMATION:  
;; APPLICANT: BEACH, David H.  
;; TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO  
;; NUMBER OF SEQUENCES: 50  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: LAHIVE & COCKFIELD  
;; STREET: 60 State Street  
;; CITY: Boston  
;; STATE: MA  
;; COUNTRY: USA  
;; ZIP: 02109  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: ASCII(text)  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/464,517  
;; FILING DATE:  
;; CLASSIFICATION: 435  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/963,308  
;; FILING DATE: 16-OCT-1992  
;; INFORMATION FOR SEQ ID NO: 7:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 173 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single

;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/701,514  
;; FILING DATE: 16-MAY-1991  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Matthew P. Vincent  
;; REGISTRATION NUMBER: 36,709  
;; REFERENCE/DOCKET NUMBER: MII-004C  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (617) 227-7400  
;; TELEFAX: (617) 227-5941  
;; INFORMATION FOR SEQ ID NO: 21:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 189 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
;; US-08-464-517-21

Query Match 61.5%; Score 32; DB 2; Length 189;  
Best Local Similarity 60.0%; Pred. No. 42;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10  
||| | | |  
DB 74 EEVFPPLAMNY 83

RESULT 17  
US-08-246-361A-21  
;; Sequence 21, Application US/08246361A  
;; Patent No. 5998582  
;; GENERAL INFORMATION:  
;; APPLICANT: BEACH, David H.  
;; TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO  
;; NUMBER OF SEQUENCES: 50  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: LAHIVE & COCKFIELD  
;; STREET: 60 State Street  
;; CITY: Boston  
;; STATE: MA  
;; COUNTRY: USA  
;; ZIP: 02109  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: ASCII(text)  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/246,361A  
;; FILING DATE: 19-MAY-1994  
;; CLASSIFICATION: 435  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/963,308  
;; FILING DATE: 16-OCT-1992  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/888,178  
;; FILING DATE: 26-MAY-1992  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/701,514  
;; FILING DATE: 16-MAY-1991  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Matthew P. Vincent  
;; REGISTRATION NUMBER: 36,709  
;; REFERENCE/DOCKET NUMBER: MII-004C  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (617) 227-7400  
;; TELEFAX: (617) 227-5941  
;; INFORMATION FOR SEQ ID NO: 21:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 189 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single

; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-246-361A-21

Query Match 61.5%; Score 32; DB 2; Length 189;  
Best Local Similarity 60.0%; Pred. No. 42;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10  
||| | | |  
Db 74 EEVFPFLAMNY 83

## RESULT 18

US-08-463-772-21

; Sequence 21, Application US/08463772

; Patent No. 6066501

; GENERAL INFORMATION:

; APPLICANT: BEACH, David H.

; TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO

; NUMBER OF SEQUENCES: 50

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: LAHIVE & COCKFIELD

; STREET: 60 State Street

; CITY: Boston

; STATE: MA

; COUNTRY: USA

; ZIP: 02109

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: ASCII(text)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/463,772

; FILING DATE:

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/963,308

; FILING DATE: 16-OCT-1992

; APPLICATION NUMBER: US 07/888,178

; FILING DATE: 26-MAY-1992

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/701,514

; FILING DATE: 16-MAY-1991

; ATTORNEY/AGENT INFORMATION:

; NAME: Matthew P. Vincent

; REGISTRATION NUMBER: 36,709

; REFERENCE/DOCKET NUMBER: MII-004C

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (617) 227-7400

; TELEFAX: (617) 227-5941

; INFORMATION FOR SEQ ID NO: 21:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 189 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

US-08-463-772-21

Query Match 61.5%; Score 32; DB 3; Length 189;  
Best Local Similarity 60.0%; Pred. No. 42;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10  
||| | | |  
Db 74 EEVFPFLAMNY 83

## RESULT 19

PCT-US93-05000-21

; Sequence 21, Application PC/TUS9305000

; GENERAL INFORMATION:  
; APPLICANT: MITOTIX  
; TITLE OF INVENTION: D-Type Cyclin and Uses Related Thereto  
; NUMBER OF SEQUENCES: 42

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.

; STREET: Two Militia Drive

; CITY: Lexington

; STATE: Massachusetts

; COUNTRY: US

; ZIP: 02173

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: PCT/US93/05000

; FILING DATE: 19930525

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/07/888,178

; FILING DATE: 26-MAY-1992

; ATTORNEY/AGENT INFORMATION:

; NAME: Granahan, Patricia

; REGISTRATION NUMBER: 32,227

; REFERENCE/DOCKET NUMBER: CSHL91-02A

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 617-861-6240

; TELEFAX: 616-861-9540

; INFORMATION FOR SEQ ID NO: 21:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 189 amino acids

; TYPE: AMINO ACID

; TOPOLOGY: unknown

; MOLECULE TYPE: protein

PCT-US93-05000-21

Query Match 61.5%; Score 32; DB 5; Length 189;  
Best Local Similarity 60.0%; Pred. No. 42;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10  
||| | | |  
Db 74 EEVFPFLAMNY 83

## RESULT 20

US-08-464-517-22

; Sequence 22, Application US/08464517

; Patent No. 5869640

; GENERAL INFORMATION:

; APPLICANT: BEACH, David H.

; TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO

; NUMBER OF SEQUENCES: 50

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: LAHIVE & COCKFIELD

; STREET: 60 State Street

; CITY: Boston

; STATE: MA

; COUNTRY: USA

; ZIP: 02109

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: ASCII(text)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/464,517

; FILING DATE:

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/963,308

;; FILING DATE: 16-OCT-1992  
;; APPLICATION NUMBER: US 07/888,178  
;; FILING DATE: 26-MAY-1992  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/701,514  
;; FILING DATE: 16-MAY-1991  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Matthew P. Vincent  
;; REGISTRATION NUMBER: 36,709  
;; REFERENCE/DOCKET NUMBER: MII-004C  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (617) 227-7400  
;; TELEFAX: (617) 227-5941  
;; INFORMATION FOR SEQ ID NO: 22:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 236 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
US-08-464-517-22

Query Match 61.5%; Score 32; DB 2; Length 236;  
Best Local Similarity 60.0%; Pred. No. 54;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10  
||| | | |  
DB 20 EEVFPPLAMNY 29

## RESULT 21

US-08-246-361A-22  
;; Sequence 22, Application US/08246361A  
;; Patent No. 5998582  
;; GENERAL INFORMATION:  
;; APPLICANT: BEACH, David H.  
;; TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO  
;; NUMBER OF SEQUENCES: 50  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: LAHIVE & COCKFIELD  
;; STREET: 60 State Street  
;; CITY: Boston  
;; STATE: MA  
;; COUNTRY: USA  
;; ZIP: 02109  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: ASCII(text)  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/246,361A  
;; FILING DATE: 19-MAY-1994  
;; CLASSIFICATION: 435  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/963,308  
;; FILING DATE: 16-OCT-1992  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/888,178  
;; FILING DATE: 26-MAY-1992  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/701,514  
;; FILING DATE: 16-MAY-1991  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Matthew P. Vincent  
;; REGISTRATION NUMBER: 36,709  
;; REFERENCE/DOCKET NUMBER: MII-004C  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (617) 227-7400  
;; TELEFAX: (617) 227-5941  
;; INFORMATION FOR SEQ ID NO: 22:  
;; SEQUENCE CHARACTERISTICS:

;; LENGTH: 236 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
US-08-246-361A-22

Query Match 61.5%; Score 32; DB 2; Length 236;  
Best Local Similarity 60.0%; Pred. No. 54;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10  
||| | | |  
DB 20 EEVFPPLAMNY 29

## RESULT 22

US-08-463-772-22  
;; Sequence 22, Application US/08463772  
;; Patent No. 6066501  
;; GENERAL INFORMATION:  
;; APPLICANT: BEACH, David H.  
;; TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO  
;; NUMBER OF SEQUENCES: 50  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: LAHIVE & COCKFIELD  
;; STREET: 60 State Street  
;; CITY: Boston  
;; STATE: MA  
;; COUNTRY: USA  
;; ZIP: 02109  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: ASCII(text)  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/463,772  
;; FILING DATE:  
;; CLASSIFICATION: 435  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/963,308  
;; FILING DATE: 16-OCT-1992  
;; APPLICATION NUMBER: US 07/888,178  
;; FILING DATE: 26-MAY-1992  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/701,514  
;; FILING DATE: 16-MAY-1991  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Matthew P. Vincent  
;; REGISTRATION NUMBER: 36,709  
;; REFERENCE/DOCKET NUMBER: MII-004C  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (617) 227-7400  
;; TELEFAX: (617) 227-5941  
;; INFORMATION FOR SEQ ID NO: 22:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 236 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
US-08-463-772-22

Query Match 61.5%; Score 32; DB 3; Length 236;  
Best Local Similarity 60.0%; Pred. No. 54;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10  
||| | | |  
DB 20 EEVFPPLAMNY 29

RESULT 23  
PCT-US93-05000-22  
; Sequence 22, Application PC/TUS9305000  
; GENERAL INFORMATION:  
; APPLICANT: MITOTIX  
; TITLE OF INVENTION: D-Type Cyclin and Uses Related Thereto  
; NUMBER OF SEQUENCES: 42  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
; STREET: Two Militia Drive  
; CITY: Lexington  
; STATE: Massachusetts  
; COUNTRY: US  
; ZIP: 02173  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US93/05000  
; FILING DATE: 19930525  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/07/888,178  
; FILING DATE: 26-MAY-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Granahan, Patricia  
; REGISTRATION NUMBER: 32,227  
; REFERENCE/DOCKET NUMBER: CSHL91-02A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-861-6240  
; TELEFAX: 616-861-9540  
; INFORMATION FOR SEQ ID NO: 22:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 236 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
PCT-US93-05000-22

Query Match 61.5%; Score 32; DB 5; Length 236;  
Best Local Similarity 60.0%; Pred. No. 54;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10  
Db 20 EEVFPPLAMNY 29

RESULT 24  
US-08-464-517-6  
; Sequence 6, Application US/08464517  
; Patent No. 5869640  
; GENERAL INFORMATION:  
; APPLICANT: BEACH, David H.  
; TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO  
; NUMBER OF SEQUENCES: 50  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD  
; STREET: 60 State Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII(text)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/464,517  
; FILING DATE:

; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/963,308  
; FILING DATE: 16-OCT-1992  
; APPLICATION NUMBER: US 07/888,178  
; FILING DATE: 26-MAY-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/701,514  
; FILING DATE: 16-MAY-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Matthew P. Vincent  
; REGISTRATION NUMBER: 36,709  
; REFERENCE/DOCKET NUMBER: MII-004C  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 227-7400  
; TELEFAX: (617) 227-5941  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 280 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-464-517-6

Query Match 61.5%; Score 32; DB 2; Length 280;  
Best Local Similarity 60.0%; Pred. No. 65;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10  
Db 75 EEVFPPLAMNY 84

RESULT 25  
US-08-463-772-6  
; Sequence 6, Application US/08463772  
; Patent No. 6066501  
; GENERAL INFORMATION:  
; APPLICANT: BEACH, David H.  
; TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO  
; NUMBER OF SEQUENCES: 50  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD  
; STREET: 60 State Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII(text)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/463,772  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/963,308  
; FILING DATE: 16-OCT-1992  
; APPLICATION NUMBER: US 07/888,178  
; FILING DATE: 26-MAY-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/701,514  
; FILING DATE: 16-MAY-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Matthew P. Vincent  
; REGISTRATION NUMBER: 36,709  
; REFERENCE/DOCKET NUMBER: MII-004C  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 227-7400  
; TELEFAX: (617) 227-5941  
; INFORMATION FOR SEQ ID NO: 6:

; SEQUENCE CHARACTERISTICS:  
; LENGTH: 280 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-463-772-6

Query Match 61.5%; Score 32; DB 3; Length 280;  
Best Local Similarity 60.0%; Pred. No. 65;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVPEXGMSY 10  
||| | | |  
DB 75 EEVPLAMNY 84

Search completed: June 10, 2003, 13:51:31  
Job time : 9.64286 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2003, 13:46:50 ; Search time 15 Seconds  
(without alignments)  
75.710 Million cell updates/sec

Title: US-09-909-164-6  
Perfect score: 52  
Sequence: 1 BEVFXGMSYS 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 392085 seqs, 103240269 residues

Total number of hits satisfying chosen parameters: 392085

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 75 summaries

Database : Published Applications\_AA:  
1: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB pep.\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB pep.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB pep.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB pep.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB pep.\*  
6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB pep.\*  
7: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB pep.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB pep.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB pep.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB pep.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB pep.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB pep.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB pep.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	38	73.1	3472	9	US-10-027-806-4
2	38	73.1	3472	9	US-10-034-623-4
3	38	73.1	3472	9	US-10-027-801-4
4	34	65.4	947	9	US-10-101-464A-73
5	33	63.5	426	9	US-10-214-766-43
6	33	63.5	478	9	US-09-924-340-108
7	33	63.5	478	9	US-09-992-600A-108
8	33	63.5	478	9	US-09-746-783-184
9	33	63.5	478	9	US-10-000-489-108
10	33	63.5	478	9	US-10-000-986-108
11	33	63.5	653	9	US-09-820-843A-26
12	32	61.5	10	10	US-09-947-387-66
13	32	61.5	254	10	US-09-778-927A-53
14	32	61.5	289	9	US-10-024-066-2
15	32	61.5	289	9	US-10-024-066-4
16	32	61.5	289	10	US-09-919-497-54
17	32	61.5	295	10	US-09-925-300-1061
18	32	61.5	529	10	US-09-923-304-4
19	32	61.5	691	9	US-10-101-921-4

20	61.5	691	10	US-09-925-731-2	Sequence 2, Appli
21	1377	1377	9	US-09-815-242-10384	Sequence 10384, A
22	61.5	2799	9	US-10-151-736-4	Sequence 4, Appli
23	61.5	53	9	US-10-092-154-878	Sequence 878, App
24	59.6	53	10	US-09-764-847-878	Sequence 878, App
25	59.6	59	10	US-09-948-080-14	Sequence 14, Appl
26	59.6	161	9	US-09-738-626-5124	Sequence 20, Appl
27	59.6	163	9	US-10-117-846-20	Sequence 171, App
28	59.6	192	9	US-09-986-480-171	Sequence 334, App
29	59.6	198	10	US-09-731-872-334	Sequence 6349, App
30	59.6	223	9	US-09-738-626-6349	Sequence 10697, A
31	59.6	299	10	US-09-815-242-10697	Sequence 2, Appli
32	59.6	381	9	US-09-975-139-5	Sequence 2, Appli
33	59.6	702	9	US-10-280-403-2	Sequence 454, Ap
34	59.6	702	10	US-09-907-479-2	Sequence 454, Ap
35	59.6	763	9	US-09-738-626-4454	Sequence 5136, Ap
36	59.6	1053	10	US-09-815-242-5136	Sequence 10439, A
37	59.6	1407	10	US-09-815-242-10439	Sequence 340, App
38	59.6	1426	10	US-09-912-020-340	Sequence 1, Appli
39	57.7	7	9	US-09-909-062-1	Sequence 9, Appli
40	57.7	7	9	US-09-909-062-9	Sequence 130, App
41	57.7	7	9	US-09-909-062-130	Sequence 68, Appl
42	57.7	121	9	US-09-852-797-68	Sequence 85, Appl
43	57.7	121	9	US-09-852-797-85	Sequence 85, Appl
44	57.7	121	10	US-09-853-161-68	Sequence 85, Appl
45	57.7	121	10	US-09-853-161-85	Sequence 85, Appl
46	57.7	121	10	US-09-852-659A-68	Sequence 85, Appl
47	57.7	121	10	US-09-852-659A-85	Sequence 85, Appl
48	57.7	135	9	US-09-992-598-359	Sequence 359, App
49	57.7	135	9	US-09-989-293A-359	Sequence 359, App
50	57.7	135	9	US-09-989-735-359	Sequence 359, App
51	57.7	135	9	US-09-990-444-359	Sequence 359, App
52	57.7	135	9	US-09-989-730-359	Sequence 359, App
53	57.7	135	9	US-09-990-436-359	Sequence 359, App
54	57.7	135	9	US-09-991-181-359	Sequence 359, App
55	57.7	135	9	US-09-993-687-359	Sequence 359, App
56	57.7	135	9	US-09-989-734-359	Sequence 359, App
57	57.7	135	9	US-09-997-653-359	Sequence 359, App
58	57.7	135	9	US-10-174-590-444	Sequence 444, App
59	57.7	135	9	US-10-176-758-444	Sequence 444, App
60	57.7	135	9	US-10-175-737-444	Sequence 444, App
61	57.7	135	9	US-09-993-667-359	Sequence 359, App
62	57.7	135	9	US-10-173-706-444	Sequence 444, App
63	57.7	135	9	US-10-175-738-444	Sequence 444, App
64	57.7	135	9	US-10-175-752-444	Sequence 444, App
65	57.7	135	9	US-10-176-482-444	Sequence 444, App
66	57.7	135	9	US-10-176-757-444	Sequence 444, App
67	57.7	135	9	US-10-176-913-444	Sequence 444, App
68	57.7	135	9	US-10-180-552-444	Sequence 444, App
69	57.7	135	9	US-10-180-557-444	Sequence 444, App
70	57.7	135	9	US-09-990-438-359	Sequence 359, App
71	57.7	135	9	US-09-990-562-359	Sequence 359, App
72	57.7	135	9	US-09-997-428-359	Sequence 359, App
73	57.7	135	9	US-09-997-666-359	Sequence 359, App
74	57.7	135	9	US-10-173-700-444	Sequence 444, App
75	57.7	135	9	US-10-174-572-444	Sequence 444, App

## ALIGNMENTS

## RESULT 1

US-10-027-806-4  
; Sequence 4, Application US/10027806  
; Patent No. US20020160476A1  
; GENERAL INFORMATION:  
; APPLICANT: Swanson, Ronald V.  
; APPLICANT: Feldman, Robert A.  
; APPLICANT: Schleper, Christa  
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM CENARCHAEUM SYMBIOSUM  
; FILE REFERENCE: DCCORP.002A  
; CURRENT APPLICATION NUMBER: US/10/027,806  
; CURRENT FILING DATE: 2001-12-21

; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/408,020  
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-09-29  
; NUMBER OF SEQ ID NOS: 123  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 4  
; LENGTH: 3472  
; TYPE: PRT  
; ORGANISM: Cenarchaeum symbiosum  
US-10-027-806-4

Query Match 73.1%; Score 38; DB 9; Length 3472;  
Best Local Similarity 54.5%; Pred. No. 1.2e+02;  
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXGMSYS 11  
|:|:| |:|:|  
Db 2294 EDVIPRGISFS 2304

## RESULT 2

US-10-034-623-4  
; Sequence 4, Application US/10034623  
; Publication No. US20020198365A1  
; GENERAL INFORMATION:  
; APPLICANT: Swanson, Ronald V.  
; APPLICANT: Feldman, Robert A.  
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM CENARCHAEUM SYMBIOSUM  
; FILE REFERENCE: DCCP.002A  
; CURRENT APPLICATION NUMBER: US/10/034,623  
; CURRENT FILING DATE: 2001-12-21  
; PRIOR APPLICATION NUMBER: 09/408,020  
; PRIOR FILING DATE: 1999-09-29  
; PRIOR APPLICATION NUMBER: 60/102,294  
; PRIOR FILING DATE: 1998-09-29  
; NUMBER OF SEQ ID NOS: 123  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 4  
; LENGTH: 3472  
; TYPE: PRT  
; ORGANISM: Cenarchaeum symbiosum  
US-10-034-623-4

Query Match 73.1%; Score 38; DB 9; Length 3472;  
Best Local Similarity 54.5%; Pred. No. 1.2e+02;  
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXGMSYS 11  
|:|:| |:|:|  
Db 2294 EDVIPRGISFS 2304

## RESULT 3

US-10-027-801-4  
; Sequence 4, Application US/10027801  
; Publication No. US20030054364A1  
; GENERAL INFORMATION:  
; APPLICANT: Swanson, Ronald V.  
; APPLICANT: Feldman, Robert A.  
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM CENARCHAEUM SYMBIOSUM  
; FILE REFERENCE: DCCP.002A  
; CURRENT APPLICATION NUMBER: US/10/027,801  
; CURRENT FILING DATE: 2001-12-21  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/408,020  
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-09-29  
; NUMBER OF SEQ ID NOS: 123  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 4  
; LENGTH: 3472  
; TYPE: PRT  
; ORGANISM: Cenarchaeum symbiosum  
US-10-027-801-4

Query Match 73.1%; Score 38; DB 9; Length 3472;  
Best Local Similarity 54.5%; Pred. No. 1.2e+02;  
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXGMSYS 11

Query Match 73.1%; Score 38; DB 9; Length 3472;  
Best Local Similarity 54.5%; Pred. No. 1.2e+02;  
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXGMSYS 11  
|:|:| |:|:|  
Db 2294 EDVIPRGISFS 2304

## RESULT 4

US-10-101-464A-73  
; Sequence 73, Application US/10101464A  
; Publication No. US20030046728A1  
; GENERAL INFORMATION:  
; APPLICANT: Strabala, Timothy  
; APPLICANT: Nieuwenhuizen, Nicolaas  
; APPLICANT: Higgins, Colleen M.  
; TITLE OF INVENTION: Compositions Isolated from Plant Cells  
; FILE REFERENCE: 11000.1020c2  
; CURRENT APPLICATION NUMBER: US/10/101,464A  
; CURRENT FILING DATE: 2002-03-18  
; PRIOR APPLICATION NUMBER: 09/704,302  
; PRIOR FILING DATE: 2000-11-01  
; PRIOR APPLICATION NUMBER: 09/228,986  
; PRIOR FILING DATE: 1999-01-12  
; PRIOR APPLICATION NUMBER: 60/162,866  
; PRIOR FILING DATE: 1999-11-01  
; PRIOR APPLICATION NUMBER: PCT/US00/00724  
; PRIOR FILING DATE: 2000-01-11  
; NUMBER OF SEQ ID NOS: 989  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 73  
; LENGTH: 947  
; TYPE: PRT  
; ORGANISM: Pinus radiata  
US-10-101-464A-73

Query Match 65.4%; Score 34; DB 9; Length 947;  
Best Local Similarity 66.7%; Pred. No. 1.8e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMSYS 11  
|:|:| |:|:|  
Db 686 VMPXGMSYS 694

## RESULT 5

US-10-214-766-43  
; Sequence 43, Application US/10214766  
; Publication No. US20030084473A1  
; GENERAL INFORMATION:  
; APPLICANT: Gocal, Greg  
; TITLE OF INVENTION: NON-TRANSGENIC HERBICIDE RESISTANT PLANTS  
; FILE REFERENCE: CA1138  
; CURRENT APPLICATION NUMBER: US/10/214,766  
; CURRENT FILING DATE: 2002-08-09  
; PRIOR APPLICATION NUMBER: US 60/311,734  
; PRIOR FILING DATE: 2001-08-09  
; NUMBER OF SEQ ID NOS: 53  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 43  
; LENGTH: 426  
; TYPE: PRT  
; ORGANISM: Vibrio cholerae  
US-10-214-766-43

Query Match 63.5%; Score 33; DB 9; Length 426;  
Best Local Similarity 60.0%; Pred. No. 1.2e+02;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSYS 10



Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 EVVPXGMSYS 11  
Db 239 EVAPAGASYN 248

RESULT 8  
US-09-746-783-184  
; Sequence 184, Application US/09746783  
; Publication No. US20030044935A1  
; GENERAL INFORMATION:  
; APPLICANT: Jacobs, Kenneth  
; McCoy, John M.  
; Lavallie, Edward R.  
; Racie, Lisa A.  
; Treacy, Maurice  
; Spaulding, Vikki  
; Agostino, Michael J.  
; Howes, Steven H.  
; Fechtel, Kim  
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES  
; NUMBER OF SEQUENCES: 231  
; CORRESPONDENCE ADDRESS:  
; STREET: 87 CambridgePark Drive  
; CITY: Cambridge  
; STATE: MA  
; COUNTRY: U.S.A.  
; ZIP: 02140  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/746,783  
; FILING DATE: 21-Dec-2000  
; CLASSIFICATION: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Milasincic, Debra J.  
; REGISTRATION NUMBER: 46,931  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 227-7400  
; TELEFAX: (617) 742-4214  
; INFORMATION FOR SEQ ID NO: 184:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 478 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: <Unknown>  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 184:  
US-09-746-783-184

Query Match 63.5%; Score 33; DB 9; Length 478;  
Best Local Similarity 60.0%; Pred. No. 1.4e+02;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 EVVPXGMSYS 11  
Db 239 EVAPAGASYN 248

RESULT 9  
US-10-000-489-108  
; Sequence 108, Application US/10000489  
; Publication No. US20030092011A1  
; GENERAL INFORMATION:  
; APPLICANT: Benjamin, Stephane  
; APPLICANT: Tanaka, Hiroaki  
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF

Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 EVVPXGMSYS 11  
Db 239 EVAPAGASYN 248

RESULT 6  
US-09-924-340-108  
; Sequence 108, Application US/09924340  
; Publication No. US20030027248A1  
; GENERAL INFORMATION:  
; APPLICANT: Benjamin, Stephane  
; APPLICANT: Tanaka, Hiroaki  
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF  
; FILE REFERENCE: 91.052.REG  
; CURRENT APPLICATION NUMBER: US/09/924,340  
; CURRENT FILING DATE: 2001-08-06  
; PRIOR APPLICATION NUMBER: US 60/305,456  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: US 60/302,277  
; PRIOR FILING DATE: 2001-06-29  
; PRIOR APPLICATION NUMBER: US 60/298,698  
; PRIOR FILING DATE: 2001-06-15  
; PRIOR APPLICATION NUMBER: US 60/293,574  
; PRIOR FILING DATE: 2001-05-25  
; NUMBER OF SEQ ID NOS: 112  
; SOFTWARE: JPatent  
; SEQ ID NO 108  
; LENGTH: 478  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-924-340-108

Query Match 63.5%; Score 33; DB 9; Length 478;  
Best Local Similarity 60.0%; Pred. No. 1.4e+02;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 EVVPXGMSYS 11  
Db 239 EVAPAGASYN 248

RESULT 7  
US-09-992-600A-108  
; Sequence 108, Application US/09992600A  
; Publication No. US20030027161A1  
; GENERAL INFORMATION:  
; APPLICANT: Benjamin, Stephane  
; APPLICANT: Tanaka, Hiroaki  
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF  
; FILE REFERENCE: 91.054.DIV  
; CURRENT APPLICATION NUMBER: US/09/992,600A  
; CURRENT FILING DATE: 2001-11-13  
; PRIOR APPLICATION NUMBER: US 09/924,340  
; PRIOR FILING DATE: 2001-08-06  
; PRIOR APPLICATION NUMBER: PCT/IB01/01715  
; PRIOR FILING DATE: 2001-08-06  
; PRIOR APPLICATION NUMBER: US 60/305,456  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: US 60/302,277  
; PRIOR FILING DATE: 2001-06-29  
; PRIOR APPLICATION NUMBER: US 60/298,698  
; PRIOR FILING DATE: 2001-06-15  
; PRIOR APPLICATION NUMBER: US 60/293,574  
; PRIOR FILING DATE: 2001-05-25  
; NUMBER OF SEQ ID NOS: 114  
; SOFTWARE: JPatent  
; SEQ ID NO 108  
; LENGTH: 478  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-992-600A-108

Query Match 63.5%; Score 33; DB 9; Length 478;  
Best Local Similarity 60.0%; Pred. No. 1.4e+02;

; FILE REFERENCE: 91.US6.DIV  
; CURRENT APPLICATION NUMBER: US/10/000,489  
; CURRENT FILING DATE: 2001-11-14  
; PRIOR APPLICATION NUMBER: US 09/924,340  
; PRIOR FILING DATE: 2001-08-06  
; PRIOR APPLICATION NUMBER: PCT/IB01/01715  
; PRIOR FILING DATE: 2001-08-06  
; PRIOR APPLICATION NUMBER: US 60/305,456  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: US 60/302,277  
; PRIOR FILING DATE: 2001-06-29  
; PRIOR APPLICATION NUMBER: US 60/298,698  
; PRIOR FILING DATE: 2001-06-15  
; PRIOR APPLICATION NUMBER: US 60/293,574  
; PRIOR FILING DATE: 2001-05-25  
; NUMBER OF SEQ ID NOS: 112  
; SOFTWARE: JPpatent  
; SEQ ID NO 108  
; LENGTH: 478  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-000-489-108

Query Match 63.5%; Score 33; DB 9; Length 478;  
Best Local Similarity 60.0%; Pred. No. 1.4e+02;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 EVVPXGMSYS 11  
||| |||  
Db 239 EVAPAGASYN 248

RESULT 10  
US-10-000-986-108  
; Sequence 108, Application US/10000986  
; Publication No. US20030096247A1  
; GENERAL INFORMATION:  
; APPLICANT: Benjamin, Stephanie  
; APPLICANT: Tanaka, Hiroaki  
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF  
; FILE REFERENCE: 91.US9.DIV  
; CURRENT APPLICATION NUMBER: US/10/000,986  
; CURRENT FILING DATE: 2001-11-14  
; PRIOR APPLICATION NUMBER: US 09/924,340  
; PRIOR FILING DATE: 2001-08-06  
; PRIOR APPLICATION NUMBER: PCT/IB01/01715  
; PRIOR FILING DATE: 2001-08-06  
; PRIOR APPLICATION NUMBER: US 60/305,456  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: US 60/302,277  
; PRIOR FILING DATE: 2001-06-29  
; PRIOR APPLICATION NUMBER: US 60/298,698  
; PRIOR FILING DATE: 2001-06-15  
; PRIOR APPLICATION NUMBER: US 60/293,574  
; PRIOR FILING DATE: 2001-05-25  
; NUMBER OF SEQ ID NOS: 112  
; SOFTWARE: JPpatent  
; SEQ ID NO 108  
; LENGTH: 478  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-000-986-108

Query Match 63.5%; Score 33; DB 9; Length 478;  
Best Local Similarity 60.0%; Pred. No. 1.4e+02;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 EVVPXGMSYS 11  
||| |||  
Db 239 EVAPAGASYN 248

RESULT 11

US-09-820-843A-26  
; Sequence 26, Application US/09820843A  
; Publication No. US20030039963A1  
; GENERAL INFORMATION:  
; APPLICANT: Council of Scientific and Industrial Research  
; TITLE OF INVENTION: A COMPUTATIONAL METHOD FOR THE IDENTIFICATION OF CANDIDATE  
; FILE REFERENCE: Q63915  
; CURRENT APPLICATION NUMBER: US/09/820,843A  
; CURRENT FILING DATE: 2001-03-30  
; NUMBER OF SEQ ID NOS: 118  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 26  
; LENGTH: 653  
; TYPE: PRT  
; ORGANISM: Vibrio cholerae  
; FEATURE:  
; NAME/KEY: misc.feature  
; OTHER INFORMATION: iron(III) ABC transporter, permease protein  
; NAME/KEY: misc.feature  
; OTHER INFORMATION: gi|9654609  
US-09-820-843A-26

Query Match 63.5%; Score 33; DB 9; Length 653;  
Best Local Similarity 66.7%; Pred. No. 1.9e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXGMS 9  
||||| |||  
Db 300 EEVVPXGMS 308

RESULT 12  
US-09-947-387-66  
; Sequence 66, Application US/09947387  
; Patent No. US20020150885A1  
; GENERAL INFORMATION:  
; APPLICANT: Weber, Eckard  
; APPLICANT: Cai, Sui Xiong  
; APPLICANT: Keana, John F.W.  
; APPLICANT: Drewe, John A.  
; APPLICANT: Zhang, Han-Zhong  
; TITLE OF INVENTION: No. US20020150885A1el Fluorogenic or Fluorescent Reporter Mol  
; TITLE OF INVENTION: Their Applications for Whole-Cell Fluorescence  
; TITLE OF INVENTION: Screening Assays for Caspases and Other Enzymes and the  
; FILE REFERENCE: 1735.0290005  
; CURRENT APPLICATION NUMBER: US/09/947,387  
; CURRENT FILING DATE: 2001-09-07  
; PRIOR APPLICATION NUMBER: US 60/061,582  
; PRIOR FILING DATE: 1997-10-10  
; PRIOR APPLICATION NUMBER: US 60/145,746  
; PRIOR FILING DATE: 1998-03-03  
; PRIOR APPLICATION NUMBER: US 09/168,888  
; PRIOR FILING DATE: 1998-10-09  
; NUMBER OF SEQ ID NOS: 142  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 66  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
; OTHER INFORMATION: Peptide  
US-09-947-387-66

Query Match 61.5%; Score 32; DB 10; Length 10;  
Best Local Similarity 50.0%; Pred. No. 3.5;  
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10  
::: |||  
Db 1 DDIVPCMSY 10

## RESULT 13

US-09-978-927A-53

; Sequence 53, Application US/09778927A

; Patent No. US20020068342A1

; GENERAL INFORMATION:

; APPLICANT: KHOSRAVI, Rami et al.

; TITLE OF INVENTION: NOVEL NUCLEIC ACID AND AMINO ACID SEQUENCES AND NOVEL

; TITLE OF INVENTION: VARIANTS OF ALTERNATIVE SPLICING

; FILE REFERENCE: 2786-0160P

; CURRENT APPLICATION NUMBER: US/09/778,927A

; CURRENT FILING DATE: 2001-02-08

; PRIOR APPLICATION NUMBER: IL 134453

; PRIOR FILING DATE: 2000-02-09

; PRIOR APPLICATION NUMBER: IL135341

; PRIOR FILING DATE: 2000-03-29

; NUMBER OF SEQ ID NOS: 81

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 53

; LENGTH: 254

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc\_feature

; LOCATION: (1)..(254)

; OTHER INFORMATION: Xaa = any amino acid, unknown, or other

US-09-778-927A-53

## Query Match

Best Local Similarity 61.5%; Score 32; DB 10; Length 254;

Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy

1 EEVVPXGMSY 10

|||||

Db 74 EEVFPPLMNY 83

## RESULT 14

US-10-024-066-2

; Sequence 2, Application US/10024066

; Patent No. US20020166134A1

; GENERAL INFORMATION:

; APPLICANT: Field, Loren J.

; TITLE OF INVENTION: CARDIOMYOCYTES WITH ENHANCED PROLIFERATIVE POTENTIAL,

; TITLE OF INVENTION: AND METHODS FOR PREPARING AND USING SAME

; FILE REFERENCE: 7037-450

; CURRENT APPLICATION NUMBER: US/10/024,066

; CURRENT FILING DATE: 2001-12-18

; PRIOR APPLICATION NUMBER: 60/139,942

; PRIOR FILING DATE: 1999-06-18

; PRIOR APPLICATION NUMBER: PCT/US00/16827

; PRIOR FILING DATE: 2000-06-19

; NUMBER OF SEQ ID NOS: 8

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 2

; LENGTH: 289

; TYPE: PRT

; ORGANISM: Mus musculus

US-10-024-066-2

## Query Match

Best Local Similarity 61.5%; Score 32; DB 9; Length 289;

Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy

1 EEVVPXGMSY 10

|||||

Db 73 EEVFPPLMNY 82

## RESULT 15

US-10-024-066-4

; Sequence 4, Application US/10024066

; Patent No. US20020166134A1

; GENERAL INFORMATION:

; APPLICANT: Field, Loren J.

; APPLICANT: Pasumarthi, Kishore Babu S.

; TITLE OF INVENTION: CARDIOMYOCYTES WITH ENHANCED PROLIFERATIVE POTENTIAL,

; TITLE OF INVENTION: AND METHODS FOR PREPARING AND USING SAME

; FILE REFERENCE: 7037-450

; CURRENT APPLICATION NUMBER: US/10/024,066

; CURRENT FILING DATE: 2001-12-18

; PRIOR APPLICATION NUMBER: 60/139,942

; PRIOR FILING DATE: 1999-06-18

; PRIOR APPLICATION NUMBER: PCT/US00/16827

; PRIOR FILING DATE: 2000-06-19

; NUMBER OF SEQ ID NOS: 8

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 4

; LENGTH: 289

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-024-066-4

## Query Match

Best Local Similarity 61.5%; Score 32; DB 9; Length 289;

Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy

1 EEVVPXGMSY 10

|||||

Db 74 EEVFPPLMNY 83

## RESULT 16

US-09-919-497-54

; Sequence 54, Application US/09919497

; Patent No. US20020106662A1

; GENERAL INFORMATION:

; APPLICANT: Mutter, George L.

; TITLE OF INVENTION: PROGNOSTIC CLASSIFICATION OF ENDOMETRIAL CANCER

; FILE REFERENCE: B0801/7225

; CURRENT APPLICATION NUMBER: US/09/919,497

; CURRENT FILING DATE: 2001-07-31

; PRIOR APPLICATION NUMBER: US 60/221,735

; PRIOR FILING DATE: 2000-07-31

; NUMBER OF SEQ ID NOS: 100

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 54

; LENGTH: 289

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-919-497-54

## Query Match

Best Local Similarity 61.5%; Score 32; DB 10; Length 289;

Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy

1 EEVVPXGMSY 10

|||||

Db 74 EEVFPPLMNY 83

## RESULT 17

US-09-925-300-1061

; Sequence 1061, Application US/09925300

; Patent No. US20020151681A1

; GENERAL INFORMATION:

; APPLICANT: Craig Rosen,

; APPLICANT: Steve Ruben

; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies

; FILE REFERENCE: PA101

; CURRENT APPLICATION NUMBER: US/09/925,300

; CURRENT FILING DATE: 2001-08-10

; PRIOR APPLICATION NUMBER: PCT/US00/05988

; PRIOR FILING DATE: 2000-03-08

```

; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1061
; LENGTH: 295
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (243)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (277)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-300-1061

Query Match      61.5%; Score 32; DB 10; Length 295;
Best Local Similarity 50.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      2 EVVPXGMSYS 11
      ||| |||
Db      52 EVLPTKMSYA 61

RESULT 18
US-09-923-304-4
; Sequence 4, Application US/09923304
; Patent No. US20020081612A1
; GENERAL INFORMATION:
; APPLICANT: KATZ, RUTH
; APPLICANT: JIANG, FENG
; TITLE OF INVENTION: DETECTION AND DIAGNOSIS OF SMOKING RELATED CANCERS
; FILE REFERENCE: UTSC:638US
; CURRENT APPLICATION NUMBER: US/09/923,304
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 529
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-923-304-4

Query Match      61.5%; Score 32; DB 10; Length 529;
Best Local Similarity 50.0%; Pred. No. 2.4e+02;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY      1 BEVPXGMSY 10
      ||| |||
Db      48 QEVLPPLGLKY 57

RESULT 19
US-10-101-921-4
; Sequence 4, Application US/10101921
; Publication No. US20030022199A1
; GENERAL INFORMATION:
; APPLICANT: Nezu, Jun-Ichi
; APPLICANT: Ose, Asuka
; APPLICANT: Tsuji, Akira
; TITLE OF INVENTION: TRANSPORTER GENES OATP-B, C, D, AND E
; FILE REFERENCE: 06501-104US1
; CURRENT APPLICATION NUMBER: US/10/101,921
; CURRENT FILING DATE: 2002-06-28
; PRIOR APPLICATION NUMBER: PCT/JP00/06416
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: JP 11/267835
; PRIOR FILING DATE: 1999-09-21
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4

```

```

; LENGTH: 691
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-101-921-4

Query Match      61.5%; Score 32; DB 9; Length 691;
Best Local Similarity 62.5%; Pred. No. 3.2e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      3 VVPXGMSY 10
      ||| |||
Db      188 IVPLGLSY 195

RESULT 20
US-09-925-731-2
; Sequence 2, Application US/09925731
; Patent No. US20020090622A1
; GENERAL INFORMATION:
; APPLICANT: ADEOKUN, ANTHONI MONISOLA
; APPLICANT: AMBROSE, HELEN JEAN
; APPLICANT: CRESSWELL, CARL JOHN
; APPLICANT: DUDLEY, ADAM JESTON
; TITLE OF INVENTION: CHEMICAL COMPOUNDS
; FILE REFERENCE: DJB/009901/0282795
; CURRENT APPLICATION NUMBER: US/09/925,731
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/226,909
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 691
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-731-2

Query Match      61.5%; Score 32; DB 10; Length 691;
Best Local Similarity 62.5%; Pred. No. 3.2e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      3 VVPXGMSY 10
      ||| |||
Db      188 IVPLGLSY 195

RESULT 21
US-09-815-242-10384
; Sequence 10384, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625

```

```
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10384
; LENGTH: 1377
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-815-242-10384

Query Match      61.5%; Score 32; DB 10; Length 1377;
Best Local Similarity 45.5%; Pred. No. 6.7e+02;
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY      1 EEVVPXGMSYS 11
Db      369 EQLNPAGLSYT 379
      ||::|||::|

RESULT 22
US-10-151-736-4
; Sequence 4, Application US/10151736
; Publication No. US20020192160A1
; GENERAL INFORMATION:
; APPLICANT: Callaghan, Michelle J.
; APPLICANT: Sutherland, Lindfield
; APPLICANT: Watts, Colin K.
; TITLE OF INVENTION: No. US20020192160A1e1 Human Tumour Suppressor Gene
; FILE REFERENCE: RICE-0100CON
; CURRENT APPLICATION NUMBER: US/10/151,736
; PRIOR FILING DATE: 2002-05-15
; PRIOR APPLICATION NUMBER: 09/403,402
; PRIOR FILING DATE: 1999-10-21
; PRIOR APPLICATION NUMBER: PCT/AU98/00280
; PRIOR FILING DATE: 1998-04-20
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 2799
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-151-736-4

Query Match      61.5%; Score 32; DB 9; Length 2799;
Best Local Similarity 60.0%; Pred. No. 1.4e+03;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      2 EEVVPXGMSYS 11
Db      2096 EVLPTRMSYA 2105
      ||::|||::|

RESULT 23
US-10-092-154-878
; Sequence 878, Application US/10092154
; Publication No. US20030054375A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC009C1
; CURRENT APPLICATION NUMBER: US/10/092,154
; CURRENT FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 2003
; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 878
; LENGTH: 53
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:

; NAME/KEY: misc_feature
; LOCATION: (7)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-092-154-878

Query Match      59.6%; Score 31; DB 9; Length 53;
Best Local Similarity 66.7%; Pred. No. 33;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      3 VVPXGMSYS 11
Db      39 VVPTAVSYS 47
      |||::|||

RESULT 24
US-09-764-847-878
; Sequence 878, Application US/09764847
; Patent No. US20020132767A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC009
; CURRENT APPLICATION NUMBER: US/09/764,847
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2003
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 878
; LENGTH: 53
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (7)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-847-878

Query Match      59.6%; Score 31; DB 10; Length 53;
Best Local Similarity 66.7%; Pred. No. 33;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      3 VVPXGMSYS 11
Db      39 VVPTAVSYS 47
      |||::|||

RESULT 25
US-09-948-080-14
; Sequence 14, Application US/09948080
; Patent No. US20020102702A1
; GENERAL INFORMATION:
; APPLICANT: VAN DER OSTEN, CLAUS
; APPLICANT: HALKIER, TORDEN
; APPLICANT: ANDERSEN, CARSTEN
; APPLICANT: BAUDITZ, PETER
; APPLICANT: HANSEN, PETER KAMP
; TITLE OF INVENTION: PROTEASE VARIANTS AND COMPOSITIONS
; FILE REFERENCE: 4946,200-US
; CURRENT APPLICATION NUMBER: US/09/948,080
; CURRENT FILING DATE: 2001-09-06
; PRIOR APPLICATION NUMBER: US/08/963,851
; PRIOR FILING DATE: 1997-11-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 14
; LENGTH: 59
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-948-080-14

Query Match      59.6%; Score 31; DB 10; Length 59;
Best Local Similarity 45.5%; Pred. No. 37;
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
```

Wed Jun 11 15:44:17 2003

us-09-909-164-6.rapb

Page 8

Qy 1 EEVVPXGMSYS 11  
J: : : : :  
Db 38 EKHIPGGLYS 48

Search completed: June 10, 2003, 14:35:40  
Job time : 15.0714 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2003, 13:31:15 ; Search time 11.2143 Seconds  
(without alignments)  
94.297 Million cell updates/sec

Title: US-09-909-164-6  
Perfect score: 52  
Sequence: 1 EEVVPXGMSYS 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues  
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 75 summaries

Database : PIR\_73:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	38	73.1	3472	T31308	hypothetical 367K
2	37	71.2	840	T39116	probable sulfate p
3	37	71.2	877	T40413	sulfate permease -
4	36	69.2	102	A42452	V1 protein - tobac
5	36	69.2	1498	B97355	DNA segregation At
6	35	67.3	225	S57810	hypothetical prote
7	35	67.3	425	T24111	hypothetical prote
8	35	67.3	670	S22293	zinc finger protei
9	35	67.3	749	H82691	topoisomerase IV s
10	35	67.3	2717	A34203	DNA-binding protei
11	34	65.4	156	S54619	hypothetical prote
12	34	65.4	252	H69491	cell division inhi
13	34	65.4	544	C82900	probable ABC subst
14	33	63.5	94	T40758	hypothetical prote
15	33	63.5	116	E90544	50S ribosomal prot
16	33	63.5	165	D69493	hypothetical prote
17	33	63.5	253	C81374	hypothetical prote
18	33	63.5	259	T34536	hypothetical prote
19	33	63.5	284	S75817	hypothetical prote
20	33	63.5	298	T47670	beta-ketoacyl-ACP
21	33	63.5	368	F72281	hypothetical prote
22	33	63.5	426	D82163	3-phosphoshikimate
23	33	63.5	466	T43653	cdc37 protein - fi
24	33	63.5	653	D82352	iron(III) ABC tran
25	33	63.5	890	A30481	bacteriocin BCN5 -
26	33	63.5	1028	Ar3286	ATP-dependent DNA
27	33	63.5	1152	D87046	conserved hypothet
28	33	63.5	1394	A35626	transforming growt
29	33	63.5	1401	G82336	DNA-directed RNA p

30	63.5	1548	2	T04456	hypothetical prote
31	63.5	1712	2	A38261	masking protein pr
32	61.5	84	2	E97333	hypothetical prote
33	61.5	175	2	PQ0616	transport protein
34	61.5	223	2	T01457	rho protein GDP-di
35	61.5	279	2	B72481	hypothetical prote
36	61.5	288	2	JC4011	cyclin D2 - rat
37	61.5	288	2	I58372	cyclin D2 - rat
38	61.5	289	2	A41984	cyclin D2 - mouse
39	61.5	289	2	A42822	cyclin D2 - human
40	61.5	291	2	S57922	cyclin D1 - Africa
41	61.5	291	2	S57925	cyclin D2 - Africa
42	61.5	291	2	JC4579	cyclin D2 - chicke
43	61.5	291	2	S62730	cyclin D1 - zebra
44	61.5	292	2	B42822	cyclin D3 - human
45	61.5	295	2	A38977	cyclin D1 - human
46	61.5	295	2	A56523	cyclin D1 - mouse
47	61.5	295	2	JC2342	cyclin D1 - rat
48	61.5	347	2	I55120	rhsF protein - Esc
49	61.5	363	2	D69551	conserved hypothet
50	61.5	427	2	F64064	tolB protein - Hae
51	61.5	449	2	A99286	conserved hypothet
52	61.5	498	2	B90604	agaA protein limpo
53	61.5	525	2	D98311	hypothetical prote
54	61.5	525	2	AF2971	transport protein
55	61.5	726	2	T44000	infected cell prot
56	61.5	726	2	T44187	SC11 protein - yea
57	61.5	759	2	S25330	hypothetical prote
58	61.5	889	2	S22659	maltooligosyltreha
59	61.5	922	2	AG1827	RNA 1 protein - to
60	61.5	993	1	P1VXTA	aggregation protei
61	61.5	1306	2	S22624	rhaA protein precu
62	61.5	1377	2	C65159	rhaA protein in rh
63	61.5	1377	2	E86034	RhsH core protein
64	61.5	1394	2	H91236	rhaA protein precu
65	61.5	1397	2	A85570	rhaC protein in rh
66	61.5	1397	2	C64805	RhsC core protein
67	61.5	1399	2	A99720	rhaA core protein
68	61.5	1409	2	F91187	rhaA core protein
69	61.5	1411	2	E65145	rhaA protein precu
70	61.5	1948	2	S00485	gene 11-1 protein
71	59.6	124	1	VKLJ51	trans-regulatory s
72	59.6	133	2	A71173	hypothetical prote
73	59.6	208	1	D70764	probable cobH - My
74	59.6	224	2	G70709	probable purQ prot
75	59.6	224	2	F87186	phosphoribosylform

## ALIGNMENTS

## RESULT 1

T31308  
hypothetical 367K protein - Cenarchaeum symbiosum

C:Species: Cenarchaeum symbiosum

C>Date: 11-Jan-2000 #sequence\_revision 11-Jan-2000 #text\_change 18-Feb-2000

C:Accession: T31308

R.Schleper, C.; DeLong, E.F.; Preston, C.M.; Feldman, R.A.; Wu, K.Y.; Swanson, R.V.

J. Bacteriol. 180, 5003-5009, 1998

A:Title: Genomic analysis reveals chromosomal variation in natural populations of t

A:Reference number: Z20994; MUID:98422450; PMID:9748430

A:Accession: T31308

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-3472 <SCH>

A:Cross-references: EMBL:AF083072; NID:g3599393; PID:g3599394; PIDN:AAC62699.1

C:Superfamily: Cenarchaeum symbiosum hypothetical 367K protein

Query Match 73.1%; Score 38; DB 2; Length 3472;  
Best Local Similarity 54.5%; Pred. No. 59;  
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXGMSYS 11

Db 2294 EDVIPRGISFS 2304  
|:|:| |:|:|

## RESULT 2

T39116  
probable sulfate permease - fission yeast (Schizosaccharomyces pombe)  
C:Species: Schizosaccharomyces pombe  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Dec-1999  
C:Accession: T39116  
R:Hunt, C.; Aves, S.; McDougal, R.C.; Rajandream, M.A.; Barrell, B.G.  
submitted to the EMBL Data Library, November 1999  
A:Reference number: 221829  
A:Accession: T39116  
A:Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-840 <HUN>  
A:Cross-references: EMBL:AL132779; PIDN:CAB60015.1; GSPDB:GN00066; SPDB:SPAC869.05C  
A:Experimental source: strain 972h-; cosmid c869  
C:Genetics:  
A:Gene: SPDB:SPAC869.05C  
A:Map position: 1

Query Match 71.2%; Score 37; DB 2; Length 840;

Best Local Similarity 77.8%; Pred. No. 21;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMSYS 11  
||| |||||  
Db 135 VVPGMSYA 143

## RESULT 3

T40413  
sulfate permease - fission yeast (Schizosaccharomyces pombe)  
C:Species: Schizosaccharomyces pombe  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Dec-1999  
C:Accession: T40413  
R:Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Jimenez Martinez, J.  
submitted to the EMBL Data Library, August 1998  
A:Reference number: 221926  
A:Accession: T40413  
A:Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-877 <LYN>  
A:Cross-references: EMBL:AL031261; PIDN:CAA20298.1; GSPDB:GN00067; SPDB:SPBC3H7.02  
A:Experimental source: strain 972h-; cosmid c3H7  
C:Genetics:  
A:Gene: SPDB:SPBC3H7.02  
A:Map position: 2

Query Match 71.2%; Score 37; DB 2; Length 877;

Best Local Similarity 77.8%; Pred. No. 22;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMSYS 11  
||| |||||  
Db 148 VVPGMSYA 156

## RESULT 4

A42452  
V1 protein - tobacco yellow dwarf virus (strain Australia)  
C:Species: tobacco yellow dwarf virus  
C:Date: 15-Jan-1993 #sequence\_revision 15-Jan-1993 #text\_change 08-Oct-1999  
C:Accession: A42452  
R:Morris, B.A.M.; Richardson, K.A.; Haley, A.; Zhan, X.; Thomas, J.E.  
Virology 187, 633-642, 1992  
A:Title: The nucleotide sequence of the infectious cloned DNA component of tobacco yellow  
A:Reference number: A42452; MUID:92188538; PMID:1546458  
A:Accession: A42452  
A:Molecule type: DNA  
A:Residues: 1-102 <MOR>

A:Cross-references: GB:M81103; NID:9335283; PIDN:AAA47947.1; PID:9335284

Query Match 69.2%; Score 36; DB 2; Length 102;  
Best Local Similarity 60.0%; Pred. No. 3.5;  
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 EEVVPXGMSYS 11  
:||| |:|  
Db 7 QVVPXGMSYS 16

## RESULT 5

B97355  
DNA segregation ATPase, FtsK/SpoIIIE family, YUKA B. subtilis ortholog [imported]  
C:Species: Clostridium acetobutylicum  
C:Date: 14-Sep-2001 #sequence\_revision 14-Sep-2001 #text\_change 14-Sep-2001  
C:Accession: B97355  
R:Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.;  
J. Bacteriol. 183, 4823-4838, 2001  
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium  
A:Reference number: A96900; MUID:21359325; PMID:21359325  
A:Accession: B97355  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1498 <KUR>  
A:Cross-references: GB:AE001437; PIDN:AAK81629.1; PID:g15026814; GSPDB:GN00168  
A:Experimental source: Clostridium acetobutylicum ATCC824  
C:Genetics:  
A:Gene: CAC3709

Query Match 69.2%; Score 36; DB 2; Length 1498;  
Best Local Similarity 60.0%; Pred. No. 63;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10  
|:|:| |||||  
Db 1276 EQIPMGMSY 1285

## RESULT 6

S57810  
hypothetical protein precursor (clone TPp11) - tomato  
C:Species: Lycopersicon esculentum (tomato)  
C:Date: 28-Oct-1995 #sequence\_revision 03-Nov-1995 #text\_change 21-Jul-2000  
C:Accession: S57810  
R:Milligan, S.B.; Gasser, C.S.  
Plant Mol. Biol. 28, 691-711, 1995  
A:Title: Nature and regulation of pistil-expressed genes in tomato.  
A:Reference number: S57808; MUID:95375233; PMID:7647301  
A:Accession: S57810  
A:Status: preliminary; nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1-225 <ML>  
A:Cross-references: EMBL:U20592; NID:9924625; PIDN:AAA80497.1; PID:9924626  
C:Superfamily: plant Kunitz-type proteinase inhibitor

Query Match 67.3%; Score 35; DB 2; Length 225;  
Best Local Similarity 54.5%; Pred. No. 13;  
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMSYS 11  
:||| |:|  
Db 32 DEVVPNGKTYA 42

## RESULT 7

T24111  
hypothetical protein R10D12.10 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
C:Accession: T24111  
R:Percy, C.



submitted to the EMBL Data Library, October 1996

A:Reference number: Z19842  
 A:Accession: T24111  
 A>Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-425 <WIL>  
 A:Cross-references: EMBL:Z81109; PIDN:CAB03241.1; GSPDB:GN00023; CESP:R10D12.10  
 A:Experimental source: clone R10D12  
 C:Genetics:  
 A:Gene: CESP:R10D12.10  
 A:Map position: 5  
 A:Introns: 23/3; 56/3; 113/3; 257/2

Query Match 67.3%; Score 35; DB 2; Length 425;  
 Best Local Similarity 50.0%; Pred. No. 26;  
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EVVPXGMSY 10  
 ||| |||  
 Db 335 EQIVPGLOY 344

RESULT 8  
 S22293  
 zinc finger protein AT-BP2 - rat (fragment)  
 C:Species: Rattus norvegicus (Norway rat)  
 C:Date: 29-Jan-1998 #sequence\_revision 06-Feb-1998 #text\_change 20-Sep-1999  
 C:Accession: S22293; I78656  
 R:MitcheMore, C.; Traboni, C.; Cortese, R.  
 Nucleic Acids Res. 19, 141-147, 1991  
 A:Title: Isolation of two cDNAs encoding zinc finger proteins which bind to the alpha 1-  
 A:Reference number: I58280; MUID:91187610; PMID:1901405  
 A:Accession: S22293  
 A>Status: nucleic acid sequence not shown  
 A:Molecule type: mRNA  
 A:Residues: 1-670 <MIT>  
 A:Cross-references: EMBL:X54250; NID:G57519; PIDN:CAA38151.1; PID:G57520  
 A>Note: the authors did not translate the codon for residue 1  
 C:Superfamily: HIV-EP2 enhancer-binding protein  
 C:Keywords: DNA binding; transcription regulation; zinc finger

Query Match 67.3%; Score 35; DB 2; Length 670;  
 Best Local Similarity 56.7%; Pred. No. 43;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMSYS 11  
 ||| |||  
 Db 376 VVPAGLTYS 384

RESULT 9  
 H82691  
 topoisomerase IV subunit Xfl1353 [imported] - Xylella fastidiosa (strain 9a5c)  
 C:Species: Xylella fastidiosa  
 C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Sep-2000  
 C:Accession: H82691  
 R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen  
 Nature 406, 151-157, 2000  
 A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.  
 A:Reference number: A82515; MUID:20365717; PMID:10910347  
 A>Note: for a complete list of authors see reference number A59328 below  
 A:Accession: H82691  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-749 <STM>  
 A:Cross-references: GB:AE003967; GB:AE003849; NID:G9106347; PIDN:AAF84162.1; GSPDB:GN001  
 A:Experimental source: strain 9a5c  
 R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A  
 Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H  
 as-Neto, E.; Docena, C.; El-Dorri, H.; Facincan, A.P.; Ferreira, A.J.S.  
 submitted to GenBank, June 2000  
 A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm  
 J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig

chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martini  
 A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.  
 F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri  
 Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sa  
 M.; Tshuko, M.H.; Vallada, H.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Si  
 A:Reference number: A59328  
 A:Contents: annotation  
 C:Genetics:  
 A:Gene: Xfl1353  
 C:Superfamily: DNA topoisomerase (ATP-hydrolyzing) chain A; phage T4 DNA topoisomer  
 Query Match 67.3%; Score 35; DB 2; Length 749;  
 Best Local Similarity 77.8%; Pred. No. 48;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 EVVPXGMSY 10  
 ||| |||  
 Db 526 EVDPSGMSY 534

RESULT 10  
 A34203  
 DNA-binding protein PRDII-BF1 - human  
 N:Alternate names: major histocompatibility complex enhancer-binding protein 1  
 C:Species: Homo sapiens (man)  
 C:Date: 22-Jun-1990 #sequence\_revision 22-Jun-1990 #text\_change 20-Sep-1999  
 C:Accession: A34203; A34779  
 R:Fan, C.M.; Maniatis, T.  
 Genes Dev. 4, 29-42, 1990  
 A:Title: A DNA-binding protein containing two widely separated zinc finger motifs  
 A:Reference number: A34203; MUID:90169514; PMID:2106471  
 A:Accession: A34203  
 A>Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-2717 <FAN>  
 A:Cross-references: EMBL:X51435; NID:G38017; PIDN:CAA35798.1; PID:G38018  
 R:Balwin Jr., A.S.; LeClair, K.P.; Singh, H.; Sharp, P.A.  
 Mol. Cell. Biol. 10, 1406-1414, 1990  
 A:Title: A large protein containing zinc finger domains binds to related sequence  
 A:Reference number: A34779; MUID:90205817; PMID:2108316  
 A:Accession: A34779  
 A>Status: preliminary; nucleic acid sequence not shown  
 A:Molecule type: mRNA  
 A:Residues: 801-1072, 'N', 1074-1168, 'K', 1170-1225, 'V', 1227-1434, 'N', 1436-1607, 'I', 16  
 A:Cross-references: GB:M32019  
 C:Superfamily: HIV-EP2 enhancer-binding protein  
 C:Keywords: DNA binding; transcription regulation; zinc finger

Query Match 67.3%; Score 35; DB 2; Length 2717;  
 Best Local Similarity 66.7%; Pred. No. 1.9e+02;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMSYS 11  
 ||| |||  
 Db 2405 VVPAGLTYS 2413

RESULT 11  
 S54619  
 hypothetical protein YOR013w - yeast (Saccharomyces cerevisiae)  
 N:Alternate names: hypothetical protein O2612; hypothetical protein YOL303.3  
 C:Species: Saccharomyces cerevisiae  
 C:Date: 08-Jul-1995 #sequence\_revision 01-Sep-1995 #text\_change 19-Apr-2002  
 C:Accession: S54619; S66879  
 R:de Haan, M.; Maarse, A.C.; Grivell, L.A.  
 submitted to the EMBL Data Library, May 1995  
 A:Reference number: S54617  
 A:Accession: S54619  
 A:Molecule type: DNA  
 A:Residues: 1-156 <DEH>  
 A:Cross-references: EMBL:X87331; NID:G1041652; PIDN:CAA60762.1; PID:G829123  
 R:de Haan, M.; Grivell, L.A.; Maarse, A.C.

submitted to the Protein Sequence Database, July 1996

A:Reference number: S66877  
 A:Accession: S66879  
 A:Molecule type: DNA  
 A:Residues: 1-156 <DEW>  
 A:Cross-references: EMBL:274920; NID:g1420109; PIDN:CAA99201.1; PID:g1420111; MIPS:YOR0101  
 A:Experimental source: strain S288C  
 C:Genetics:  
 A:Cross-references: SGD:S0005539  
 A:Map position: 15R  
 C:Superfamily: hypothetical protein YOR013w

Query Match 65.4%; Score 34; DB 2; Length 156;  
 Best Local Similarity 66.7%; Pred. No. 14;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 EVVPXGMSY 10  
 ||| |||  
 Db 50 EVMLGMDY 58

# RESULT 12

H69491  
 cell division inhibitor (mind-2) homolog - Archaeoglobus fulgidus  
 C:Species: Archaeoglobus fulgidus  
 C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 26-Aug-1999  
 C:Accession: H69491  
 R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson  
 ; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.  
 Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.  
 Nature 390, 364-370, 1997  
 A:Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.  
 Smith, H.O.; Woese, C.R.; Venter, J.C.  
 A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archae  
 A:Reference number: A69250; MUID:98049343; PMID:9389475  
 A:Accession: H69491  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-252 <LE>  
 A:Cross-references: GB:AE000970; GB:AE000782; NID:g2689293; PIDN:AAB89318.1; PID:g364860  
 C:Superfamily: cell division inhibitor mind

Query Match 65.4%; Score 34; DB 2; Length 252;  
 Best Local Similarity 75.0%; Pred. No. 24;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 EVVPXGMS 9  
 ||| |||  
 Db 81 EVIPAGMS 88

# RESULT 13

C82900  
 probable ABC substrate-binding protein, iron UU359 [imported] - Ureaplasma urealyticum  
 C:Species: Ureaplasma urealyticum  
 C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Sep-2000  
 C:Accession: C82900  
 R:Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Cassell, G.H.  
 Submitted to GenBank, February 2000  
 A:Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a min  
 A:Reference number: A82870  
 A:Accession: C82900  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-544 <GLA>  
 A:Cross-references: GB:AE002133; GB:AF222894; NID:g6899339; PIDN:AAF30768.1; GSPDB:GN001  
 A:Experimental source: serovar 3; biovar 1  
 C:Genetics:  
 A:Gene: ABCsbp-5; UU359  
 A:Genetic code: SGC3

Query Match 65.4%; Score 34; DB 2; Length 544;  
 Best Local Similarity 70.0%; Pred. No. 55;

Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 EVVPXGMSY 10  
 ||||| |||  
 Db 135 EVVPHYLSY 144

# RESULT 14

I40758  
 hypothetical protein 1 - Campylobacter jejuni (fragment)  
 C:Species: Campylobacter jejuni  
 C:Date: 16-Aug-1996 #sequence\_revision 16-Aug-1996 #text\_change 08-Oct-1999  
 C:Accession: I40758; S47317  
 R:Hani, E.K.; Chan, V.L.  
 J. Bacteriol. 177, 2396-2402, 1995  
 A:Title: Expression and characterization of Campylobacter jejuni benzoylglutamine amide  
 A:Reference number: I40758; MUID:95247673; PMID:7730270  
 A:Accession: I40758  
 A:Status: preliminary; translated from GB/EMBL/DDBB  
 A:Molecule type: DNA  
 A:Residues: 1-94 <RES>  
 A:Cross-references: EMBL:Z36940; NID:g535805; PIDN:CAA85392.1; PID:g535806

Query Match 63.5%; Score 33; DB 2; Length 94;  
 Best Local Similarity 55.6%; Pred. No. 13;  
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 EVVPXGMSY 10  
 :: |||||  
 Db 26 DIFPSGMSY 34

# RESULT 15

E90544  
 50S ribosomal protein L20 [imported] - Mycoplasma pulmonis (strain UAB CTIP)  
 C:Species: Mycoplasma pulmonis  
 C:Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 03-Aug-2001  
 C:Accession: E90544  
 R:Chambaud, I.; Heilig, R.; Ferris, S.; Barbe, V.; Samson, D.; Galisson, F.; Moszer,  
 Nucleic Acids Res. 29, 2145-2153, 2001  
 A:Title: The complete genome sequence of the murine respiratory pathogen Mycoplasma  
 A:Reference number: A99512; MUID:21267165; PMID:11353084  
 A:Accession: E90544  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-116 <KUR>  
 A:Cross-references: GB:AL445566; PID:g14089674; PIDN:CAC13434.1; GSPDB:GN00153  
 A:Experimental source: strain UAB CTIP  
 C:Genetics:  
 A:Gene: MYPU\_2610  
 A:Genetic code: SGC3  
 C:Superfamily: Escherichia coli ribosomal protein L20

Query Match 63.5%; Score 33; DB 2; Length 116;  
 Best Local Similarity 77.8%; Pred. No. 17;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 VVPXGMSYS 11  
 || |||||  
 Db 68 VRPLGMSYS 76

# RESULT 16

D69493  
 hypothetical protein AF1949 - Archaeoglobus fulgidus  
 C:Species: Archaeoglobus fulgidus  
 C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 04-Mar-2000  
 C:Accession: D69493  
 R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Do  
 ; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness,  
 Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.  
 Nature 390, 364-370, 1997  
 A:Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes

Smith, H.O.; Woese, C.R.; Venter, J.C.  
 A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo  
 A:Reference number: A69250; MUID:98049343; PMID:9389475

A:Accession: D69493  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-165 <KLE>  
 A:Cross-references: GB:AE000968; GB:AE000782; NID:g2689291; PIDN:AAB89307.1; PID:g264859  
 C:Superfamily: Archaeoglobus fulgidus hypothetical protein AF1949

Query Match 63.5%; Score 33; DB 2; Length 165;  
 Best Local Similarity 60.0%; Pred. No. 25;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EVVPXGMSY 10  
 II : I I I I  
 DB 60 EESIPDGASY 69

## RESULT 17

C81374  
 hypothetical protein Cj0990c [imported] - Campylobacter jejuni (strain NCTC 11168)  
 C:Species: Campylobacter jejuni  
 C:Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 03-Jun-2002  
 C:Accession: C81374  
 R:Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chilling  
 C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Barrall  
 Nature 403, 665-668, 2000  
 A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hyp  
 A:Reference number: A81250; MUID:20150912; PMID:10688204

A:Accession: C81374  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-253 <PAR>  
 A:Cross-references: GB:AL111168; NID:g6968128; PIDN:CAB73246.1; PID:g696842  
 A:Experimental source: serotype O2, strain NCTC 11168  
 C:Genetics:  
 A:Gene: Cj0990c

Query Match 63.5%; Score 33; DB 2; Length 253;  
 Best Local Similarity 55.6%; Pred. No. 39;  
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 EVVPXGMSY 10  
 : : I I I I  
 DB 185 DIPFGMSY 193

## RESULT 18

T34536  
 hypothetical protein DKFzp34C031.1 - human (fragment)  
 C:Species: Homo sapiens (man)  
 C:Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 29-Oct-1999  
 C:Accession: T34536  
 R:Poustka, A.; Wellenreuther, R.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.  
 submitted to the Protein Sequence Database, October 1999

A:Reference number: 221540  
 A:Accession: T34536  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-259 <POU>  
 A:Cross-references: EMBL:AL122063  
 A:Experimental source: adult testis; clone DKFzp434C031  
 C:Genetics:  
 A:Note: DKFzp34C031.1

Query Match 63.5%; Score 33; DB 2; Length 259;  
 Best Local Similarity 60.0%; Pred. No. 40;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 EVVPXGMSY 11  
 I I I I I I  
 DB 22 EVAPAGASYN 31

## RESULT 19

S75817  
 hypothetical protein slr1275 - Synechocystis sp. (strain PCC 6803)  
 C:Species: Synechocystis sp.  
 A:Variety: PCC 6803  
 C:Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 08-Oct-1999  
 C:Accession: S75817  
 R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima,  
 O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Ya  
 DNA Res. 3, 109-136, 1996  
 A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechoc

S:Reference number: S74322; MUID:97061201; PMID:8905231

A:Accession: S75817  
 A:Status: nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-284 <KAN>  
 A:Cross-references: EMBL:D90913; GB:AB001339; NID:g1653348; PIDN:BAA18276.1; PID:d1  
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 63.5%; Score 33; DB 2; Length 284;  
 Best Local Similarity 55.6%; Pred. No. 44;  
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMSYS 11  
 I : I I I I  
 DB 208 VIPAGVSYT 216

## RESULT 20

T47670  
 beta-ketoacyl-ACP reductase-like protein - Arabidopsis thaliana  
 N:Alternate names: protein T26112.190  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 18-Aug-2000  
 C:Accession: T47670  
 R:Monfort, A.; Casacuberta, E.; Puigdomenech, P.; Mewes, H.W.; Lemcke, K.; Mayer, K.  
 submitted to the Protein Sequence Database, February 2000

A:Reference number: 224471  
 A:Accession: T47670  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-298 <MON>  
 A:Cross-references: EMBL:AL132954  
 A:Experimental source: cultivar Columbia; BAC clone T26112  
 C:Genetics:  
 A:Map position: 3  
 A:Introns: 25/3  
 A:Note: T26112.190  
 C:Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology

Query Match 63.5%; Score 33; DB 2; Length 298;  
 Best Local Similarity 55.6%; Pred. No. 47;  
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMSYS 11  
 : I I : I I  
 DB 187 IVPGGLAYS 195

## RESULT 21

F72281  
 hypothetical protein TM1216 - Thermotoga maritima (strain MSB8)  
 C:Species: Thermotoga maritima  
 C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 21-Jul-2000  
 C:Accession: F72281  
 R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.;  
 Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson  
 C.M.

Nature 399, 323-329, 1999

A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from geno

A:Reference number: A72200; MUID:99287316; PMID:10360571

A:Accession: F72281

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-368 <ARN>

A:Cross-references: GB:AE001778; GB:AE000512; NID:g4981757; PIDN:AAD36291.1; PID:g498178

A:Experimental source: strain MSB8

C:Genetics:

A:Gene: TM1216

C:Superfamily: NADH dehydrogenase (ubiquinone) 49K protein

Query Match

Best Local Similarity 63.5%; Score 33; DB 2; Length 368;

Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMSYS 11

Db :||| |||:

294 IVPKGMAYA 302

RESULT 22

D82163

C:Species: Vibrio cholerae

C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Feb-2001

C:Accession: D82163

R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;

chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, E.

l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000

A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

A:Reference number: A82035; MUID:20406833; PMID:10952301

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-426 <HEI>

A:Cross-references: GB:AE004251; GB:AE003852; NID:g9656248; PIDN:AAF94882.1; GSPDB:GN001

A:Experimental source: serogroup O1; strain N16961; biotype El Tor

C:Genetics:

A:Gene: VC1732

A:Map position: 1

C:Superfamily: 3-phosphoshikimate 1-carboxyvinyltransferase; 3-phosphoshikimate 1-carboxy

Query Match

Best Local Similarity 63.5%; Score 33; DB 2; Length 426;

Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10

Db :||| |||:

223 EFVIPAGQSY 232

RESULT 23

T43653

C:Species: Schizosaccharomyces pombe

C:Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 18-Feb-2000

C:Accession: T43653; T40791; T43654

R:Westwood, P.K.; Preston, N.C.; Fantes, P.A.

submitted to the EMBL Data Library, March 1999

A:Description: Schizosaccharomyces pombe cdc37 cDNA.

A:Reference number: 222602

A:Accession: T43653

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-466 <WES>

A:Cross-references: EMBL:AJ132376; PIDN:CAB38757.1

R:Wood, V.; Rajandream, M.A.; Barrell, B.G.; Volckaert, G.

submitted to the EMBL Data Library, May 1999

A:Reference number: 221875

A:Accession: T40791

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-466 <WOO>

A:Cross-references: EMBL:AL049769; PIDN:CAB42371.2; GSPDB:GN00067

A:Experimental source: strain 972h-; cosmid c9B6

R:Westwood, P.K.; Preston, N.C.; Fantes, P.A.

submitted to the EMBL Data Library, March 1999

A:Description: Schizosaccharomyces pombe cdc37 gene.

A:Reference number: 222603

A:Accession: T43654

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-466 <WE2>

A:Cross-references: EMBL:AJ132377; PIDN:CAB38758.1

C:Genetics:

A:Gene: cdc37; SPAC9B6.10

A:Map position: 2

A:Introns: 8/2; 17/2; 21/1

Query Match

Best Local Similarity 63.5%; Score 33; DB 2; Length 466;

Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10

Db :||| |||:

98 DSAIPGMSY 107

RESULT 24

D82352

C:Species: Vibrio cholerae

C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Feb-2001

C:Accession: D82352

R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson,

chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers,

l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000

A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

A:Reference number: A82035; MUID:20406833; PMID:10952301

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-653 <HEI>

A:Cross-references: GB:AE004110; GB:AE003852; NID:g9654600; PIDN:AAF93379.1; GSPDB:GN001

A:Experimental source: serogroup O1; strain N16961; biotype El Tor

C:Genetics:

A:Gene: VC0203

A:Map position: 1

Query Match

Best Local Similarity 63.5%; Score 33; DB 2; Length 653;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXGMS 9

Db :||| |||:

300 EEVVPSGIT 308

RESULT 25

A30481

C:Species: Clostridium perfringens

C:Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 24-Nov-1999

C:Accession: A30481; S03779

R:Garnier, T.; Cole, S.T.

J. Bacteriol. 168, 1189-1196, 1986

A:Title: Characterization of a bacteriocinogenic plasmid from Clostridium perfringens

A:Reference number: JT0354; MUID:87057020; PMID:2877971

A:Accession: A30481

A:Molecule type: DNA

A:Residues: 1-890 <GAR>

A:Cross-references: GB:M32882; GB:J03309; NID:g150738; PIDN:AAA98249.1; PID:g150739

C:Genetics:

A:Gene: bcn

Wed Jun 11 15:44:22 2003

us-09-909-164-6.rpr

A:Genome: plasmid  
C:Superfamily: Clostridium perfringens plasmid pIP404 bacteriocin BCN5  
C:Keywords: bacteriocin

Query Match 63.5%; Score 33; DB 2; Length 890;  
Best Local Similarity 66.7%; Pred No. 1.5e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 EVVPXGMSY 10  
||| | : |  
Db 170 EVVPGGFTY 178

Search completed: June 10, 2003, 13:49:09  
Job time : 11.2143 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2003, 13:25:04 ; Search time 4.5 Seconds  
(without alignments)  
101.387 Million cell updates/sec

Title: US-09-909-164-6  
Perfect score: 52  
Sequence: 1 BEVVPXGMSYS 11

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 75 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	38	73.1	1058	1 CARB_FUSNN	Q8rg86 fusobacteri
2	37	71.2	877	1 SULH_SCHPO	O74377 schizosacch
3	36	69.2	102	1 Y1LK_TYDVA	P31619 tobacco yel
4	36	69.2	1498	1 Y1A9_CLOAB	Q04351 clostridium
5	35	67.3	2717	1 ZEP1_HUMAN	P15822 homo sapien
6	34	65.4	788	1 CY14_NEUCR	P23622 neurospora
7	34	65.4	1499	1 A10C_HUMAN	O60312 homo sapien
8	33	63.5	116	1 RL20_MYCPU	Q38qvo mycoplasma
9	33	63.5	253	1 Y990_CAMJE	P45489 campylobact
10	33	63.5	426	1 AROA_VIBCH	Q9krb0 vibrio chol
11	33	63.5	466	1 CC37_SCHPO	O94740 schizosacch
12	33	63.5	478	1 GSR2_HUMAN	Q9nzm5 homo sapien
13	33	63.5	890	1 BCN5_CLOPE	P08696 clostridium
14	33	63.5	1394	1 LTBS_HUMAN	P22064 homo sapien
15	33	63.5	1401	1 RPOC_VIBCH	Q9kv29 vibrio chol
16	33	63.5	1595	1 LTBL_HUMAN	Q14766 homo sapien
17	33	63.5	1712	1 LTBL_RAT	Q00918 rattus norv
18	32	61.5	288	1 GGD2_RAT	Q04827 rattus norv
19	32	61.5	289	1 GGD2_HUMAN	P30279 homo sapien
20	32	61.5	289	1 CGD2_MOUSE	P30280 mus musculu
21	32	61.5	291	1 CGD1_BRARE	Q90459 brachydanio
22	32	61.5	291	1 CGD1_XENLA	P50755 xenopus lae
23	32	61.5	291	1 CGD2_CHICK	F49706 gallus gall
24	32	61.5	291	1 CGD2_XENLA	P53782 xenopus lae
25	32	61.5	292	1 CGD1_CHICK	P55169 gallus gall
26	32	61.5	292	1 CGD1_HUMAN	P30281 homo sapien
27	32	61.5	295	1 CGD1_HUMAN	P24385 homo sapien
28	32	61.5	295	1 CGD1_MOUSE	P25322 mus musculu
29	32	61.5	295	1 CGD1_RAT	F39948 rattus norv
30	32	61.5	427	1 TOLE_HAEIN	P44677 haemophilus
31	32	61.5	529	1 ENP3_HUMAN	O75355 homo sapien
32	32	61.5	691	1 OAT6_HUMAN	Q9y616 homo sapien
33	32	61.5	726	1 PRTP_HSV6U	P52384 human herpe

34	61.5	759	1	SCT1_YEAST	P32784 saccharomyc
35	61.5	920	1	EDD_RAT	O62671 rattus norv
36	61.5	993	1	VIA_TAV	P28931 tomato aspe
37	61.5	1377	1	RHSA_ECOLI	P16916 escherichia
38	61.5	1397	1	RHSE_ECOLI	P16918 escherichia
39	61.5	1411	1	RHSE_ECOLI	P16917 escherichia
40	61.5	2799	1	EDD_HUMAN	O95071 homo sapien
41	59.6	124	1	REV_SIVCZ	P17280 chimpanzee
42	59.6	208	1	COBH_MYCTU	O10676 mycobacteri
43	59.6	223	1	PURQ_PYRHO	O59619 pyrococcus
44	59.6	224	1	PURQ_HALN1	O9hnu2 halobacteri
45	59.6	224	1	PURQ_MYCLE	O05756 mycobacteri
46	59.6	224	1	PURQ_MYCTU	P71841 mycobacteri
47	59.6	225	1	PURQ_CORAM	O9rhx0 corynebacte
48	59.6	240	1	GDIR_ARATH	O9sf66 arabidopsis
49	59.6	276	1	Y939_METJA	O58349 methanococc
50	59.6	319	1	YHAI_CRYPA	P10941 cryphonectr
51	59.6	432	1	PURA_YEAST	P80210 saccharomyc
52	59.6	488	1	NOM2_PSEAE	O9htr0 pseudomonas
53	59.6	670	1	OATP_RAT	P46720 rattus norv
54	59.6	706	1	ADDC_HUMAN	O9uey8 homo sapien
55	59.6	706	1	ADDC_MOUSE	O9qub5 mus musculu
56	59.6	827	1	PLSB_MOUSE	O61586 mus musculu
57	59.6	828	1	PLSB_RAT	P97364 rattus norv
58	59.6	1047	1	EF3_SCHPO	O94489 schizosacch
59	59.6	1407	1	RPOC_ECOLI	P00577 escherichia
60	59.6	1426	1	RHSD_ECOLI	P16919 escherichia
61	59.6	2145	1	U520_CAEEL	O9u290 caenorhabdl
62	58.7	472	1	ET2A_XENLA	P19102 xenopus lae
63	57.7	81	1	YE47_ARCFU	O28825 archaeoglob
64	57.7	121	1	TKNK_HUMAN	O9uhf0 homo sapien
65	57.7	132	1	ATPE_AQUAE	O66903 aquifex aeo
66	57.7	146	1	ATPE_LACAC	O9rgy0 lactobacill
67	57.7	218	1	PURQ_METTH	O26270 methanobact
68	57.7	223	1	PURQ_RHIME	O92pil rhizobium m
69	57.7	223	1	RPTA_BUCAL	P57489 buchnera ap
70	57.7	230	1	PURQ_METJA	O59042 methanococc
71	57.7	232	1	SCOA_HELPU	O9zle3 helicobacte
72	57.7	232	1	SCOA_HELPU	P56006 helicobacte
73	57.7	286	1	SCA6_RAT	P28433 rattus norv
74	57.7	356	1	GBA2_USTNA	P87033 ustilago ma
75	57.7	361	1	RFBF_SALTY	P26391 salmonella

## ALIGNMENTS

RESULT 1  
CARB\_FUSNN  
ID CARB\_FUSNN STANDARD; PRT; 1058 AA.  
AC O8RG86;  
DT 15-JUN-2002 (Rel. 41, Created)  
DT 15-JUN-2002 (Rel. 41, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Carbamoyl-phosphate synthase large chain (EC 6.3.5.5) (Carbamoyl-phosphate synthetase ammonia chain).  
DE CARB OR FN0422.  
GN Fusobacterium nucleatum (subsp. nucleatum).  
OS Bacteria; Fusobacteria; Fusobacterium.  
OX NCBI\_TaxID=76856;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ATCC 25586;  
RX MEDLINE=21886394; PubMed=11899109;  
RA Kapatal V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A., Bhattacharyya A., Bartman A., Gardner W., Grechkin G., Zhu L., Vasieva O., Chu L., Kogan Y., Chaga O., Goltsman E., Bernal A., Larsen N., D'Souza M., Walunas T., Pusch G., Haselkorn R., Feinstein M., Kyrpides N., Overbeek R.;  
RA "Genome sequence and analysis of the oral bacterium Fusobacterium nucleatum strain ATCC 25586.";  
RT J. Bacteriol. 184:2005-2018(2002).  
CC -|- CATALYTIC ACTIVITY: 2 ATP + L-glutamine + CO(2) + H(2)O -> 2 ADP +

phosphate + L-glutamate + carbamoyl phosphate.  
 -1- COFACTOR: Binds three manganese ions (By similarity).  
 -1- PATHWAY: Arginine biosynthesis.  
 -1- PATHWAY: Pyrimidine biosynthesis; the small (or glutamine) chain promotes the hydrolysis of glutamine to ammonia, which is used by the large (or ammonia) chain to synthesize carbamoyl phosphate (By similarity).  
 -1- SIMILARITY: BELONGS TO THE CARB FAMILY.

-----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----

EMBL; AE010554; AAL94625.1; ALT\_INIT.  
 DR InterPro; IPR005483; CPase\_L.  
 DR InterPro; IPR005479; CPase\_L\_D2.  
 DR InterPro; IPR005480; CPase\_L\_D3.  
 DR InterPro; IPR005481; CPase\_L\_N.  
 DR InterPro; IPR004362; MGS\_like.  
 DR Pfam; PF00289; CPase\_L\_chain; 2.  
 DR Pfam; PF02786; CPase\_L\_D2; 2.  
 DR Pfam; PF02787; CPase\_L\_D3; 1.  
 DR Pfam; PF02142; MGS; 1.  
 DR PRINTS; PR00098; CPASE.  
 DR PROSITE; PS00866; CPASE\_1; 2.  
 DR PROSITE; PS00867; CPASE\_2; 2.  
 KW Arginine biosynthesis; Pyrimidine biosynthesis; Ligase; Repeat;  
 KW ATP-binding; Manganese; Complete proteome.

-----  
 FT DOMAIN 1 401 CARBOXYPHOSPHATE SYNTHETIC DOMAIN.  
 FT DOMAIN 402 546 OLIGOMERIZATION DOMAIN.  
 FT DOMAIN 547 929 CARBAMOYL PHOSPHATE SYNTHETIC DOMAIN.  
 FT DOMAIN 930 1058 ALLOSTERIC DOMAIN.  
 FT REPEAT 1 546  
 FT REPEAT 547 1058  
 FT NP\_BIND 153 210 ATP (POTENTIAL).  
 FT NP\_BIND 302 352 ATP (POTENTIAL).  
 FT METAL 284 284 MANGANESE 1 (BY SIMILARITY).  
 FT METAL 298 298 MANGANESE 1 AND 2 (BY SIMILARITY).  
 FT METAL 300 300 MANGANESE 2 (BY SIMILARITY).  
 FT METAL 820 820 MANGANESE 3 (BY SIMILARITY).  
 FT METAL 832 832 MANGANESE 3 (BY SIMILARITY).  
 SQ SEQUENCE 1058 AA; 117451 MW; ED7037AF77C1E39F CRC64;

Query Match 73.1%; Score 38; DB 1; Length 1058;  
 Best Local Similarity 60.0%; Pred. No. 6.2;  
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 EVVPXGMSYS 11  
 I::I I::I  
 Db 190 EIVPGLNYS 199

RESULT 2  
 ID SULH\_SCHPO STANDARD; PRT; 877 AA.  
 AC 074377;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Probable sulfate permease C3H7.02.  
 GN SPBC3H7.02.  
 OS Schizosaccharomyces pombe (Fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OC Schizosaccharomycetes.  
 OX NCBI\_TaxID=4896;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RC STRAIN=972;  
 RX MDLINE=21848401; PubMed=11859360;  
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,  
 RA Spours J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,  
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,  
 RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,  
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,  
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,  
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,  
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,  
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitch E.,  
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,  
 RA Skellton J., Simmonds M., Squares R., Squares S., Stevens K.,  
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
 RA Woodward J., Volkart G., Aert R., Robben J., Grymonprez B.,  
 RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,  
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,  
 RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,  
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,  
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Leilaure V., Mottier S.,  
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
 RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,  
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
 RA Cerruti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,  
 RA Supakowski G.V., Ussery D., Barrell B.G., Nurse P.;  
 RL "The genome sequence of Schizosaccharomyces pombe.";  
 CC Nature 415:871-880(2002).  
 CC -1- FUNCTION: HIGH AFFINITY UPTAKE OF SULFATE INTO THE CELL (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
 CC -1- SIMILARITY: BELONGS TO THE SLC26A FAMILY OF TRANSPORTERS.  
 CC -1- SIMILARITY: CONTAINS 1 STAS DOMAIN.

-----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----

EMBL; AL031261; CAA20298.1;  
 DR InterPro; IPR002645; STAS.  
 DR InterPro; IPR001902; Sulfate\_transp.  
 DR Pfam; PF00916; Sulfate\_transp; 1.  
 DR Pfam; PF01740; STAS; 1.  
 DR TIGRFAMs; TIGR00815; sulp; 1.  
 DR PROSITE; PS01130; SLC26A; 1.  
 DR PROSITE; PS0801; STAS; 1.  
 KW Transport; Transmembrane.  
 FT TRANSMEM 133 153 POTENTIAL.  
 FT TRANSMEM 161 181 POTENTIAL.  
 FT TRANSMEM 186 206 POTENTIAL.  
 FT TRANSMEM 221 241 POTENTIAL.  
 FT TRANSMEM 243 263 POTENTIAL.  
 FT TRANSMEM 292 312 POTENTIAL.  
 FT TRANSMEM 329 349 POTENTIAL.  
 FT TRANSMEM 384 404 POTENTIAL.  
 FT TRANSMEM 424 444 POTENTIAL.  
 FT TRANSMEM 461 481 POTENTIAL.  
 FT TRANSMEM 484 504 POTENTIAL.  
 FT TRANSMEM 518 538 POTENTIAL.  
 FT TRANSMEM 543 563 POTENTIAL.  
 FT DOMAIN 594 747 STAS.  
 SQ SEQUENCE 877 AA; 96373 MW; 56995A8493371E43 CRC64;

Query Match 71.2%; Score 37; DB 1; Length 877;  
 Best Local Similarity 77.8%; Pred. No. 8.3;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMSYS 11  
 I::I I::I I::I I::I







DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Potential phospholipid-transporting ATPase VC (EC 3.6.3.1) (ATPVC)  
 DE (Aminophospholipid translocase VC).  
 GN ATP10C OR ATPVC OR KIAA0566.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE=2125279; PubMed=11326269;  
 RA Meguro M., Kashiwagi A., Mitsuya K., Nakao M., Kondo I., Saitoh S.,  
 RA Oshimura M.;  
 RA "A novel maternally expressed gene, ATP10C, encodes a putative  
 RT aminophospholipid translocase associated with Angelman syndrome.";  
 RT Nat. Genet. 28:19-20(2001).  
 RL [2]  
 RL SEQUENCE FROM N.A.  
 RP MEDLINE=21313119; PubMed=11353404;  
 RX Hertzog L.B.K., Kim S.-J., Cook E.H. Jr., Ledbetter D.H.;  
 RA "The human aminophospholipid-transporting ATPase gene ATP10C maps  
 RT adjacent to UBE3A and exhibits similar imprinted expression.";  
 RL Am. J. Hum. Genet. 68:1501-1505(2001).  
 RN [3]  
 RN SEQUENCE OF 337-1499 FROM N.A.  
 RP TISSUE=Brain;  
 RX MEDLINE=98290545; PubMed=9628581;  
 RA Nagase T., Ishikawa K.-I., Miyajima N., Tanaka A., Kotani H.,  
 RA Nomura N., Ohara O.;  
 RA "Prediction of the coding sequences of unidentified human genes. IX.  
 RT The complete sequences of 100 new cDNA clones from brain which can  
 RT code for large proteins in vitro.";  
 RL DNA Res. 5:31-39(1998).  
 CC -|- CATALYTIC ACTIVITY: ATP + H(2)O = ADP + phosphate.  
 CC -|- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).  
 CC -|- TISSUE SPECIFICITY: Wide expression, with highest levels in  
 CC kidney, followed by lung, brain, prostate, testis, ovary, and  
 CC small intestine.  
 CC -|- DISEASE: Defects in ATP10C may be a cause of Angelman syndrome  
 CC (AS), also known as 'happy puppet syndrome'.  
 CC -|- SIMILARITY: BELONGS TO THE CATION TRANSPORT ATPASES FAMILY (E1-E2  
 CC ATPASES). SUBFAMILY IV.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AB051358; BAB47392.1; -  
 DR EMBL; AY029504; AAK33100.1; -  
 DR EMBL; AY029487; AAK33100.1; JOINED.  
 DR EMBL; AY029488; AAK33100.1; JOINED.  
 DR EMBL; AY029489; AAK33100.1; JOINED.  
 DR EMBL; AY029490; AAK33100.1; JOINED.  
 DR EMBL; AY029491; AAK33100.1; JOINED.  
 DR EMBL; AY029492; AAK33100.1; JOINED.  
 DR EMBL; AY029493; AAK33100.1; JOINED.  
 DR EMBL; AY029494; AAK33100.1; JOINED.  
 DR EMBL; AY029495; AAK33100.1; JOINED.  
 DR EMBL; AY029496; AAK33100.1; JOINED.  
 DR EMBL; AY029497; AAK33100.1; JOINED.  
 DR EMBL; AY029498; AAK33100.1; JOINED.  
 DR EMBL; AY029499; AAK33100.1; JOINED.  
 DR EMBL; AY029500; AAK33100.1; JOINED.  
 DR EMBL; AY029501; AAK33100.1; JOINED.  
 DR EMBL; AY029502; AAK33100.1; JOINED.  
 DR EMBL; AY029503; AAK33100.1; JOINED.  
 DR EMBL; AB011138; BAA25492.1; -  
 DR Genew; HGNC:13547; ATP10C.

DR MIM: 605855; -  
 DR MIM: 105830; -  
 DR InterPro: IPR001757; ATPase\_E1-E2.  
 DR InterPro: IPR001454; Hignase/hydriase.  
 DR Pfam: PF00702; Hydrolase; 1.  
 DR PRINTS: PR00119; CATAPASE.  
 DR PROSITE: PS00154; ATPASE\_E1\_E2; 1.  
 KW Hydrolase; Transmembrane; Phosphorylation; Magnesium; ATP-binding;  
 KW Multigene family.  
 FT DOMAIN 1 86 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 87 106 POTENTIAL.  
 FT DOMAIN 107 110 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 111 128 POTENTIAL.  
 FT DOMAIN 129 309 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 310 332 POTENTIAL.  
 FT DOMAIN 337 362 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 363 384 POTENTIAL.  
 FT DOMAIN 385 1087 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 1088 1108 POTENTIAL.  
 FT DOMAIN 1109 1119 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 1120 1140 POTENTIAL.  
 FT DOMAIN 1141 1170 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 1171 1192 POTENTIAL.  
 FT DOMAIN 1193 1199 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 1200 1222 POTENTIAL.  
 FT DOMAIN 1223 1228 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 1229 1249 POTENTIAL.  
 FT DOMAIN 1250 1267 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 1268 1292 POTENTIAL.  
 FT DOMAIN 1293 1499 CYTOPLASMIC (POTENTIAL).  
 FT MOD\_RES 427 427 PHOSPHORYLATION (BY SIMILARITY).  
 FT METAL 1031 1031 MAGNESIUM (BY SIMILARITY).  
 FT METAL 1035 1035 MAGNESIUM (BY SIMILARITY).  
 FT DOMAIN 467 470 POLY-GLU.  
 FT CONFLICT 388 388 Q -> R (IN REF. 3).  
 SQ SEQUENCE 1499 AA; 167687 MW; D4996A4D063A68D CRC64;  
 Query Match 65.4%; Score 34; DB 1; Length 1499;  
 Best Local Similarity 72.7%; Pred. No. 62;  
 Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 EEVVPXGMSYS 11  
 Db 469 EEVVPXGMSYS 11  
 RESULT 8  
 ID RL20\_MYCPU STANDARD; PRT; 116 AA.  
 AC Q98QV0;  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 GN 50S ribosomal protein L20.  
 GN RPLT OR MYPU\_2610.  
 OS Mycoplasma pulmonis.  
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.  
 OX NCBI\_TaxID=2107;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=UAB CTIP;  
 RX MEDLINE=21267165; PubMed=11353084;  
 RA Chandaud I., Heilig R., Ferris S., Barbe V., Samson D., Galisson F.,  
 RA Moszer I., Dybvig K., Wroblewski H., Vliari A., Rocha E.P.C.,  
 RA Blanchard A.;  
 RA "The complete genome sequence of the murine respiratory pathogen  
 RT Mycoplasma pulmonis.";  
 RL Nucleic Acids Res. 29:2145-2153(2001).  
 CC -|- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 23S RIBOSOMAL RNA AND IS  
 CC NECESSARY TO THE IN VITRO ASSEMBLY PROCESS OF THE 50S RIBOSOMAL  
 CC SUBUNIT; IT IS NOT INVOLVED IN THE PROTEIN SYNTHESIZING FUNCTIONS  
 CC OF THAT SUBUNIT (BY SIMILARITY).  
 CC -|- SIMILARITY: BELONGS TO THE L20P FAMILY OF RIBOSOMAL PROTEINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; AL445563; CAC13434.1; -  
 DR Mypulist; MYPUL2610; -  
 DR InterPro; IPR001081; Ribosomal\_L20.  
 DR Pfam; PF00453; Ribosomal\_L20; 1.  
 DR PRINTS; PR00062; RIBOSOMAL\_L20.  
 DR ProDom; PD002389; Ribosomal\_L20; 1.  
 DR TIGRFAMs; TIGR01032; rplL\_bact; 1.  
 DR PROSITE; PS00937; RIBOSOMAL\_L20; 1.  
 DR Ribosomal protein; rRNA-binding; Complete proteome.  
 SQ SEQUENCE 116 AA; 13565 MW; C59C748901B18F14 CRC64;

Query Match 63.5%; Score 33; DB 1; Length 116;  
 Best Local Similarity 77.8%; Pred. No. 7;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 VVPXGMSYS 11  
 I I I I I I I  
 Db 68 VRPLGMSYS 76

RESULT 9  
 Y990\_CAMJE STANDARD; PRT; 253 AA.  
 AC P45489; OSPNV0;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Hypothetical protein Cj0990c.  
 GN Cj0990c.  
 OS Campylobacter jejuni.  
 OC Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group;  
 OC Campylobacter.  
 OX NCBI\_TaxID=197;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NCTC 11168;  
 RA Parkhill J., Wren B.W., Mungall K., Ketley J.M., Churcher C.,  
 RA Basham D., Chillingworth T., Davies R.M., Feltwell T., Holtrold S.,  
 RA Jagers K., Karlyshev A.V., Moule S., Pallen M.J., Penn C.W.,  
 RA Quail M.A., Rajandream M.A., Rutherford K.M., van Vliet A.H.M.,  
 RA Whitehead S., Barrett B.G.;  
 RT "The genome sequence of the food-borne pathogen Campylobacter jejuni  
 RT reveals hypervariable sequences.";  
 RL Nature 403:665-668(2000).  
 RN [2]  
 RP SEQUENCE OF 160-253 FROM N.A.  
 RC STRAIN=ATCC 33431 / TGH 9011;  
 RA MEDLINE=95247673; PubMed=7730270;  
 RA Hani E.K., Chan V.L.;  
 RT "Expression and characterization of Campylobacter jejuni  
 RT benzoylglycine amidohydrolase (Hippuricase) gene in Escherichia  
 coli.";  
 RL J. Bacteriol. 177:2396-2402(1995).  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; AL139076; CAB73246.1; -

DR EMBL; 236940; CAA85392.1; -  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 253 AA; 29783 MW; F96D3FF3265F8A6A CRC64;

Query Match 63.5%; Score 33; DB 1; Length 253;  
 Best Local Similarity 55.6%; Pred. No. 16;  
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 EVVFXGMSY 10  
 I I I I I I I  
 Db 185 DIFFSGMSY 193

RESULT 10  
 AROA\_VIBCH STANDARD; PRT; 426 AA.  
 ID AROA\_VIBCH  
 AC O9KRB0;  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE 3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) (5-  
 DE enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) (EPSPS).  
 GN AROA OR VC1732.  
 OS Vibrio cholerae.  
 OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.  
 OX NCBI\_TaxID=666;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=El Tor N16961 / Serotype O1;  
 RX MEDLINE=20406833; PubMed=10952301;  
 RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,  
 RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,  
 RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,  
 RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,  
 RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,  
 RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,  
 RA Fraser C.M.;  
 RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio  
 RT cholerae.";  
 RL Nature 406:477-483(2000).  
 CC -1- CATALYTIC ACTIVITY: Phosphoenolpyruvate + 3-phosphoshikimate -  
 CC phosphate + 5-O-(1-carboxyvinyl)-3-phosphoshikimate.  
 CC -1- PATHWAY: Aromatic amino acids biosynthesis; shikimate pathway;  
 CC sixth step.  
 CC -1- SUBUNIT: Monomer (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).  
 CC -1- SIMILARITY: BELONGS TO THE EPSP SYNTHASE FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; AF004251; AAF94882.1; -  
 DR TIGR; VC1732;  
 DR InterPro; IPR001986; EPSP synthase.  
 DR Pfam; PF00275; EPSP\_synthase; 1.  
 DR ProDom; PD001867; EPSP\_synthase; 1.  
 DR PROSITE; PS00104; EPSP\_SYNTHASE\_1; 1.  
 DR PROSITE; PS00885; EPSP\_SYNTHASE\_2; 1.  
 KW Aromatic amino acid biosynthesis; Transferase; Complete proteome.  
 SQ SEQUENCE 426 AA; 46101 MW; 38852D6483BFE1C3 CRC64;

Query Match 63.5%; Score 33; DB 1; Length 426;  
 Best Local Similarity 60.0%; Pred. No. 27;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EVVFXGMSY 10  
 I I I I I I I  
 Db 223 EFVIFAGQSY 232

Query Match 63.5%; Score 33; DB 1; Length 466;  
 Best Local Similarity 50.0%; Pred. No. 30;  
 Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10  
 DB 98 DSAIPGMSY 107

RESULT 12  
 GSR2\_HUMAN STANDARD; PRT; 478 AA.  
 ID GSR2\_HUMAN STANDARD; PRT; 478 AA.  
 AC Q9NZM5; Q9NPP1; Q9NPR4; Q9UFI2; Q9BTC6; Q9HAX6;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 15-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Glioma tumor suppressor candidate region gene 2 protein (p60).  
 GN GLTSCR2.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.  
 RX MEDLINE=20175430; PubMed=10708517;  
 RA Smith J.S., Tachibana I., Pohl U., Lee H.K., Thanaratasingam U.,  
 Portier B.P., Ueki K., Billings S., Ramaswamy S., Mohrenweiser H.W.,  
 Scheithauer B.W., Louis D.N., Jenkins R.B.;  
 RT "A transcript map of the chromosome 19q-Arm glioma tumor suppressor  
 region.";  
 RL Genomics 64:44-50(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Muscle;  
 RA Strausberg R.;  
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE OF 9-478 FROM N.A., AND SUBCELLULAR LOCATION.  
 RX MEDLINE=99214318; PubMed=10196275;  
 RA Bruni R., Fineschi B., Ogile W.O., Roizman B.;  
 RT "A novel cellular protein, p60, interacting with both herpes simplex  
 virus 1 regulatory proteins ICP22 and ICP0 is modified in a  
 cell-type-specific manner and is recruited to the nucleus after  
 infection.";  
 RL J. Virol. 73:3810-3817(1999).  
 RN [4]  
 RP SEQUENCE OF 12-478 FROM N.A.  
 RA Andreu N., Estivill X., Escarceller M., Sumoy L.;  
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
 RN [5]  
 RP SEQUENCE OF 218-477 FROM N.A.  
 RC TISSUE=Testis;  
 RA Poustka A., Klein M., Mewes H.-W., Gassenhuber J., Wiemann S.;  
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.  
 CC -!- SUBUNIT: Interacts with HSV-1 early proteins ICP22 and ICP0.  
 CC -!- SUBCELLULAR LOCATION: Nuclear.  
 CC -!- TISSUE SPECIFICITY: Expressed at high levels in heart and  
 pancreas, moderate levels in placenta, liver, skeletal muscle, and  
 kidney, and low levels in brain and lung.  
 CC -!- SIMILARITY: BELONGS TO THE GLTSCR2 FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL; AJ132377; CAB38758.1; -  
 CC EMBL; AJ132376; CAB38757.1; -  
 CC EMBL; AL049769; CAB42371.2; -  
 CC Chapterone; Cell division; Cell cycle.  
 CC SEQUENCE 466 AA; 52554 MW; 647238B34CABB3C5 CRC64;

RESULT 11  
 CC37\_SCHPO STANDARD; PRT; 466 AA.  
 AC O94740; 2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Hsp90 co-chaperone Cdc37 (Hsp90 chaperone protein kinase-targeting  
 subunit) (Cell division control protein 37).  
 GN CDC37 OR SPAC9B6.10.  
 OS Schizosaccharomyces pombe (Fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OC Schizosaccharomycetes.  
 OX NCBI\_TaxID=4896;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Westwood P.K., Preston N.C., Fantes P.A.;  
 RT "Schizosaccharomycetes pombe cdc37 gene.";  
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=972;  
 RX MEDLINE=2184401; PubMed=11859360;  
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,  
 Scouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,  
 Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,  
 Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,  
 Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,  
 Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,  
 James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,  
 Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,  
 Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,  
 Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,  
 Skelton J., Simmonds M., Squares R., Stevens K.,  
 Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
 Woodward J., Volkart G., Aert R., Robben J., Grymoprez B.,  
 Weltjens I., Vansireels E., Rieger M., Schaefer M., Mueller-Auer S.,  
 Gabel C., Fuchs M., Fritze C., Holzer E., Moestl D., Hilbert H.,  
 Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,  
 Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,  
 Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,  
 Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
 Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,  
 Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
 Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
 Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,  
 Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;  
 RT "The genome sequence of Schizosaccharomycetes pombe.";  
 RL Nature 415:871-880(2002).  
 CC -!- FUNCTION: With Hsp90 it forms a complex that binds to several  
 kinases, resulting in stabilization and promotion of their  
 activity (By similarity).  
 CC -!- SUBUNIT: Forms a complex with Hsp90. Interacts with a number of  
 kinases (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -!- SIMILARITY: BELONGS TO THE CDC37 FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL; AJ132377; CAB38758.1; -  
 CC EMBL; AJ132376; CAB38757.1; -  
 CC EMBL; AL049769; CAB42371.2; -  
 CC Chapterone; Cell division; Cell cycle.  
 CC SEQUENCE 466 AA; 52554 MW; 647238B34CABB3C5 CRC64;

DR EMBL; BC010095; AAH10095.1; -  
 DR EMBL; AF296124; CAB30413.1; -  
 DR EMBL; AL359335; CAB94786.1; -  
 DR EMBL; AL359336; CAB94787.1; -  
 DR EMBL; AL122063; CAB59242.1; -  
 DR SWISS-2DPAGE; O9NZM5; HUMAN.  
 DR Genew; HGNC:4333; GLTSCR2.  
 DR MIM; 605691; -  
 KW Nuclear protein; Polymorphism.  
 FT VARIANT 389 389  
 FT R -> Q.  
 FT /FTID=VAR\_011486.  
 FT GGS -> HEG (IN REF. 2; AAH04229).  
 FT G -> R (IN REF. 3).  
 FT RRKEQLWEKLAKQGBLREVRRAQAFLLNPSATRAKPGQD  
 FT TVERP -> SGRSSYGRSWPSRASPGGAQSPVAQPFEN  
 FT KGNPAPGHRITAA (IN REF. 3).  
 FT SDNPLDRPLVGQDEFFLE -> LNNPDKPVVWPGCLFPG  
 FT (IN REF. 3).  
 FT A -> S (IN REF. 2; AAH04229).  
 FT D -> H (IN REF. 3).  
 FT PEGNLLDRFKSFQRNMIEPRERAKFRKKYKVLVEKRAF  
 FT REIQ -> VLVTSVCRGAFPCVMTPTSLPPVPRGYGRHHGCP  
 FT WAGPVGPMPRG (IN REF. 5).  
 FT EGNILDRFKSFQRNMIEPRERAKFRKKYKVLVEKRAF  
 FT EIQL -> RGHSFETGSAFPGGI (IN REF. 3).  
 FT SEQUENCE 478 AA; 54417 MW; 7F18923E348CB52B CRC64;  
 Query Match 63.5%; Score 33; DB 1; Length 478;  
 Best Local Similarity 60.0%; Pred. No. 30;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 EVVPXGMSYS 11  
 |||||  
 DB 239 EVAPAGASYN 248

RESULT 13  
 BCN5\_CLOPE  
 ID BCN5\_CLOPE STANDARD; PRT; 890 AA.  
 AC P08696;  
 DT 01-JAN-1998 (Rel. 06, Created)  
 DT 01-JAN-1998 (Rel. 06, Last sequence update)  
 DT 01-OCT-1994 (Rel. 30, Last annotation update)  
 DE Bacteriocin BCN5.  
 GN BCN.  
 OS Clostridium perfringens.  
 OG Plasmid pIP404.  
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;  
 OC Clostridium.  
 OX NCBI\_TaxID=1502;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CPN50;  
 RX MEDLINE=88336297; PubMed=2901768;  
 RA Garnier T., Cole S.T.;  
 RT "Complete nucleotide sequence and genetic organization of the  
 RT bacteriocinogenic plasmid, pIP404, from Clostridium perfringens.";  
 RL Plasmid 19:134-150(1986).  
 [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CPN50;  
 RX MEDLINE=87057020; PubMed=2877971;  
 RA Garnier T., Cole S.T.;  
 RT "Characterization of a bacteriocinogenic plasmid from Clostridium  
 RT perfringens and molecular genetic analysis of the  
 RT bacteriocin-encoding gene.";  
 RL J. Bacteriol. 168:1189-1196(1986).  
 [3]  
 RP SEQUENCE OF 1-14 FROM N.A.  
 RC STRAIN=CPN50;  
 RX MEDLINE=89039249; PubMed=2460717;  
 RA Garnier T., Cole S.T.;  
 RT "Studies of UV-inducible promoters from Clostridium perfringens in

RT vivo and in vitro.";  
 RL Mol. Microbiol. 2:607-614(1988).  
 CC -1- FUNCTION: MAY FUNCTION AS AN IONOPHORE.  
 CC -1- INDUCTION: BY UV IRRADIATION.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M14481; AAA98248.1; -  
 DR EMBL; M32882; AAA98249.1; -  
 DR PIR; A30481; A30481.  
 DR InterPro; IPR003646; SH3\_bac.  
 DR SMART; SM00287; SH3b; 3.  
 KW Antibiotic; Bacteriocin; Plasmid.  
 FT DOMAIN 815 869 HYDROPHOBIC.  
 FT SEQUENCE 890 AA; 96699 MW; F4E5E8971C31C6C6 CRC64;  
 Query Match 63.5%; Score 33; DB 1; Length 890;  
 Best Local Similarity 66.7%; Pred. No. 58;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 EVVPXGMSY 10  
 |||||  
 DB 170 EVVPGGFTY 178

RESULT 14  
 LTBS\_HUMAN  
 ID LTBS\_HUMAN STANDARD; PRT; 1394 AA.  
 AC P22064;  
 DT 01-AUG-1991 (Rel. 19, Created)  
 DT 01-AUG-1991 (Rel. 19, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Latent transforming growth factor beta binding protein 1S precursor  
 DE (transforming growth factor beta-1 binding protein 1) (TGF-beta1-BP-  
 DE 1).  
 GN LTBP1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RC TISSUE=Fibroblast, and Platelet;  
 RX MEDLINE=90275601; PubMed=2350783;  
 RA Kanzaki T., Olofsson A., Moren A., Wernstedt C., Hellman U.,  
 RA Miyazono K., Claesson-Welsh L., Heldin C.-H.;  
 RT "TGF-beta 1 binding protein: a component of the large latent complex  
 RT of TGF-beta 1 with multiple repeat sequences.";  
 RL Cell 61:1051-1061(1990).  
 CC -1- SUBUNIT: THE LARGE LATENT COMPLEX OF TGF-BETA1 FROM PLATELETS IS  
 CC COMPOSED OF THE TGF-BETA1 MOLECULE NONCOVALENTLY ASSOCIATED WITH  
 CC A DISULFIDE-BONDED COMPLEX OF A DIMER OF THE N-TERMINAL PROPEPTIDE  
 CC OF THE TGF-BETA1 PRECURSOR AND A THIRD COMPONENT DENOTED TGF-  
 CC BETA1-BP. TGF-BETA1-BP DOES NOT BIND DIRECTLY TO ACTIVE TGF-BETA1.  
 CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; a short form (shown here) and a  
 CC long form (AC Q14766); are produced by alternative splicing.  
 CC -1- PTM: CONTAINS HYDROXYLATED ASPARAGINE RESIDUES.  
 CC -1- PTM: The N-terminus is blocked.  
 CC -1- SIMILARITY: CONTAINS 16 EGF-LIKE DOMAINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

FT	DISULFID	902	915	BY SIMILARITY
FT	DISULFID	921	933	BY SIMILARITY
FT	DISULFID	927	942	BY SIMILARITY
FT	DISULFID	944	957	BY SIMILARITY
FT	DISULFID	963	975	BY SIMILARITY
FT	DISULFID	970	984	BY SIMILARITY
FT	DISULFID	986	1000	BY SIMILARITY
FT	DISULFID	1101	1114	BY SIMILARITY
FT	DISULFID	1109	1123	BY SIMILARITY
FT	DISULFID	1125	1138	BY SIMILARITY
FT	DISULFID	1144	1155	BY SIMILARITY
FT	DISULFID	1150	1164	BY SIMILARITY
FT	DISULFID	1166	1179	BY SIMILARITY
FT	DISULFID	1298	1309	BY SIMILARITY
FT	DISULFID	1304	1318	BY SIMILARITY
FT	DISULFID	1320	1333	BY SIMILARITY
FT	DISULFID	1339	1354	BY SIMILARITY
FT	DISULFID	1349	1363	BY SIMILARITY
FT	DISULFID	1365	1378	BY SIMILARITY
FT	MOD_RES	647	647	HYDROXYLATION.
FT	MOD_RES	810	810	HYDROXYLATION.
FT	CARBOHYD	21	21	N-LINKED (GLCN
FT	CARBOHYD	52	52	N-LINKED (GLCN
FT	CARBOHYD	98	98	N-LINKED (GLCN
FT	CARBOHYD	294	294	N-LINKED (GLCN
FT	CARBOHYD	870	870	N-LINKED (GLCN
FT	CARBOHYD	923	923	N-LINKED (GLCN
FT	CARBOHYD	1039	1039	N-LINKED (GLCN
FT				/FTId-CAR_0001
SQ	SEQUENCE	1394 AA;	152791 MW;	DFFCA81A401
Query Match 63.5%; Score 33; DB				
Best Local Similarity 45.5%; Pred. No. 93;				
Matches 5; Conservative 3; Mismatches				
Qy	1	EEVVPXGMSYS	11	
		: : : : :		
Db	399	KEICPGMGYT	409	
RESULT 15				
RPQC_VIBCH				
ID	RPQC_VIBCH	STANDARD;	PRT;	1401 AA
AC	Q9KV29;			
DT	15-JUN-2002 (Rel. 41, Created)			
DT	15-JUN-2002 (Rel. 41, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	beta'-directed RNA polymerase beta' chain (EC			
DE	data' chain) (RNA polymerase beta' subunit)			
GN	RPQC OR VC0329.			
OS	Vibrio cholerae.			
OC	Bacteria; Proteobacteria; gamma subdivision			
OX	NCBI_taxid=666;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=El Tor N16961 / Serotype O1;			
RX	MEDLINE=20406833; PubMed=10952301;			
RA	Heidelberg J.F., Eisen J.A., Nelson W.C., C.			
RA	Dodson R.J., Haft D.H., Hickey E.K., Peters			
RA	Gill S.R., Nelson K.E., Read T.D., Tettelin			
RA	Ermolova M.D., Vamathevan J., Bass S., Qin			
RA	McDonald J., Utterback T., Fleischmann R.D.			
RA	Salzberg S.L., Smith H.O., Colwell R.R., Mel			
RA	Fraser C.M.;			
RT	"DNA sequence of both chromosomes of the cho			
RT	cholerae.";			
RL	Nature 406:477-483(2000).			
CC	-!- FUNCTION: DNA-DEPENDENT RNA POLYMERASE (			
CC	OF DNA INTO RNA USING THE FOUR RIBONUCLE			
CC	SUBSTRATES (By similarity).			
CC	-!- CATALYTIC ACTIVITY: N nucleoside triphos			
CC	{N} (N).			
CC	-!- SUBUNIT: THE ENZYME CONSISTS OF THE SIGH			







```

FT DISULFID 994 1005 BY SIMILARITY.
FT DISULFID 1000 1014 BY SIMILARITY.
FT DISULFID 1017 1028 BY SIMILARITY.
FT DISULFID 1034 1045 BY SIMILARITY.
FT DISULFID 1040 1054 BY SIMILARITY.
FT DISULFID 1056 1069 BY SIMILARITY.
FT DISULFID 1075 1086 BY SIMILARITY.
FT DISULFID 1081 1095 BY SIMILARITY.
FT DISULFID 1097 1110 BY SIMILARITY.
FT DISULFID 1116 1127 BY SIMILARITY.
FT DISULFID 1122 1136 BY SIMILARITY.
FT DISULFID 1138 1151 BY SIMILARITY.
FT DISULFID 1157 1169 BY SIMILARITY.
FT DISULFID 1164 1178 BY SIMILARITY.
FT DISULFID 1180 1192 BY SIMILARITY.
FT DISULFID 1198 1210 BY SIMILARITY.
FT DISULFID 1204 1219 BY SIMILARITY.
FT DISULFID 1221 1234 BY SIMILARITY.
FT DISULFID 1240 1252 BY SIMILARITY.
FT DISULFID 1246 1261 BY SIMILARITY.
FT DISULFID 1263 1276 BY SIMILARITY.
FT DISULFID 1282 1294 BY SIMILARITY.
FT DISULFID 1289 1303 BY SIMILARITY.
FT DISULFID 1305 1319 BY SIMILARITY.
FT DISULFID 1419 1432 BY SIMILARITY.
FT DISULFID 1427 1441 BY SIMILARITY.
FT DISULFID 1443 1456 BY SIMILARITY.
FT DISULFID 1462 1473 BY SIMILARITY.
FT DISULFID 1468 1482 BY SIMILARITY.
FT DISULFID 1484 1497 BY SIMILARITY.
FT DISULFID 1616 1627 BY SIMILARITY.
FT DISULFID 1622 1636 BY SIMILARITY.
FT DISULFID 1638 1651 BY SIMILARITY.
FT DISULFID 1657 1672 BY SIMILARITY.
FT DISULFID 1667 1681 BY SIMILARITY.
FT DISULFID 1683 1696 BY SIMILARITY.
FT CARBOHYD 339 370 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 370 379 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 416 416 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 612 612 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1042 1042 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1242 1242 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1357 1357 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 1712 AA; 186598 MW; 650BCERA691FD134 CRC64;

Query Match 63.5%; Score 33; DB 1; Length 1712;
Best Local Similarity 45.5%; Pred. No. 1.le+02;
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSYS 11
Db :: | | | |
718 KEICPGMGYT 728

RESULT 18
CGD2_RAT ID CGD2_RAT STANDARD; PRT; 288 AA.
AC Q04827;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE GI/S-specific cyclin D2 (Win-1 proto-oncogene).
GN CCND2 OR VIN-1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_Taxid=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93275661; PubMed=8502486;
RA Hanna Z., Jankowski M., Tremblay P., Jiang X.M., Milatovich A.,
RA Francke U., Jolicœur P.;
RT "The Vin-1 gene, identified by provirus insertional mutagenesis, is

```

```

RT the cyclin D2.";
RL Oncogene 8:1661-1666(1993).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=95011623; PubMed=7926809;
RA Hosokawa Y., Onga T., Nakashima K.;
RT "Induction of D2 and D3 cyclin-encoding genes during promotion of the
RT GI/S transition by prolactin in rat NB2 cells.";
RL Gene 147:249-252(1994).
CC -!- FUNCTION: ESSENTIAL FOR THE CONTROL OF THE CELL CYCLE AT THE G1/S
CC (START) TRANSITION.
CC -!- SUBUNIT: INTERACTS WITH THE CDK4 AND CDK6 PROTEIN KINASES TO FORM
CC A SERINE/THREONINE KINASE HOLOENZYME COMPLEX. THE CYCLIN SUBUNIT
CC IMPARTS SUBSTRATE SPECIFICITY TO THE COMPLEX.
CC -!- SIMILARITY: BELONGS TO THE CYCLIN FAMILY. CYCLIN D SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; L09752; AAA41010.1; -
DR EMBL; D16308; BAA03815.1; -
DR InterPro; IPR004366; Cyclin.
DR InterPro; IPR004367; Cyclin_Cterm.
DR Pfam; PF00134; cyclin; 1.
DR SMART; SM002984; cyclin_C; 1.
DR SMART; SM00385; CYCLIN; 1.
DR PROSITE; PS00292; CYCLINS; 1.
KW Cyclin; Cell cycle; Cell division; Multigene family; Proto-oncogene.
FT CONFLICT 68 68 E -> G (IN REF. 2).
FT CONFLICT 104 104 C -> V (IN REF. 2).
FT CONFLICT 232 232 T -> A (IN REF. 2).
SQ SEQUENCE 288 AA; 32826 MW; 4B522BF4E9835FC1 CRC64;

Query Match 61.5%; Score 32; DB 1; Length 288;
Best Local Similarity 60.0%; Pred. No. 29;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10
Db ||| | | |
73 EEVFPLAMNY 82

RESULT 19
CGD2_HUMAN ID CGD2_HUMAN STANDARD; PRT; 289 AA.
AC P30279; Q13955;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE GI/S-specific cyclin D2.
GN CCND2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92347851; PubMed=1386336;
RA Xiong Y., Menninger J., Beach D., Ward D.C.;
RT "Molecular cloning and chromosomal mapping of CCND genes encoding
RT human D-type cyclins.";
RL Genomics 13:575-584(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=93205384; PubMed=8455931;
RA Palmero I., Holder A., Sinclair A.J., Dickson C., Peters G.;
RT "Cyclins D1 and D2 are differentially expressed in human B-lymphoid
RT cell lines.";

```

RL Oncogene 8:1049-1054 (1993).  
RN [3]  
RP SEQUENCE FROM N.A.  
RA Miyajima N.;  
RL Submitted (MAR-1993) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Bone marrow;  
RA Strausberg R.;  
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RP SEQUENCE OF 1-240 FROM N.A.  
RC TISSUE=Placenta;  
RA Inaba T., Matsushima H., Valentine M., Roussel M.F., Sherr C.J.,  
Look A.T.;  
RT "Genomic organization, chromosomal localization, and independent  
expression of human cyclin D genes.";  
RL Genomics 13:565-574 (1992).  
CC -1- FUNCTION: ESSENTIAL FOR THE CONTROL OF THE CELL CYCLE AT THE G1/S  
(START) TRANSITION.  
CC -1- SUBUNIT: INTERACTS WITH THE CDK4 AND CDK6 PROTEIN KINASES TO FORM  
A SERINE/THREONINE KINASE HOLOENZYME COMPLEX. THE CYCLIN SUBUNIT  
IMPARTS SUBSTRATE SPECIFICITY TO THE COMPLEX.  
CC -1- SIMILARITY: BELONGS TO THE CYCLIN FAMILY. CYCLIN D SUBFAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; M30813; AAA51926.1; -;  
DR EMBL; X68452; CAA48493.1; -;  
DR EMBL; D13639; BAA02802.1; -;  
DR EMBL; BC010958; AAH10958.1; -;  
DR EMBL; M88083; AAA51928.1; -;  
DR EMBL; M88080; AAA51928.1; JOINED.  
DR EMBL; M88081; AAA51928.1; JOINED.  
DR EMBL; M88082; AAA51928.1; JOINED.  
DR PIR; A42822; A42822.  
DR PIR; S26580; S26580.  
DR Genew; HGNC:1583; CCND2.  
DR MIM; 123833; -;  
DR InterPro; IPR004366; Cyclin.  
DR InterPro; IPR004367; Cyclin\_Cterm.  
DR Pfam; PF00134; cyclin; 1.  
DR Pfam; PF02984; cyclin\_C; 1.  
DR SMART; SM00385; CYCLIN; 1.  
DR PROSITE; PS00292; CYCLINS; 1.  
DR Cyclin; Cell cycle; Cell division; Multigene family.  
KW CONFLICT 166 167 KL -> NV (IN REF. 5).  
FT CONFLICT 224 224 T -> H (IN REF. 5).  
SQ SEQUENCE 289 AA; 33067 MW; B4E5FEF476D76D90 CRC64;  
  
Query Match 61.5%; Score 32; DB 1; Length 289;  
Best Local Similarity 60.0%; Pred. No. 29;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
  
QY 1 EEVVPXGMSY 10  
DQ ||||| |  
Db 74 EEVFPPLAMNY 83  
  
RESULT 20  
CGD2\_MOUSE  
ID CGD2\_MOUSE STANDARD; PRT; 289 AA.  
AC P30280;  
DT 01-APR-1993 (Rel. 25, Created)  
DT 01-APR-1993 (Rel. 25, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE G1/S-specific cyclin D2.  
GN CCND2 OR CYL-2.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92196134; PubMed=1372445;  
RA Kiyokawa H., Busquets X., Powell C.T., Ngo L., Rifkin R.A.,  
Marks P.A.;  
RT "Cloning of a D-type cyclin from murine erythroleukemia cells.";  
RL Proc. Natl. Acad. Sci. U.S.A. 89:2444-2447 (1992).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91235305; PubMed=1827757;  
RA Matsushima H., Roussel M.F., Ashmun R.A., Sherr C.J.;  
RT "Colony-stimulating factor 1 regulates novel cyclins during the G1  
phase of the cell cycle.";  
RL Cell 65:701-713 (1991).  
CC -1- FUNCTION: ESSENTIAL FOR THE CONTROL OF THE CELL CYCLE AT THE G1/S  
(START) TRANSITION.  
CC -1- SUBUNIT: INTERACTS WITH THE CDK4 AND CDK6 PROTEIN KINASES TO FORM  
A SERINE/THREONINE KINASE HOLOENZYME COMPLEX. THE CYCLIN SUBUNIT  
IMPARTS SUBSTRATE SPECIFICITY TO THE COMPLEX.  
CC -1- SIMILARITY: BELONGS TO THE CYCLIN FAMILY. CYCLIN D SUBFAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; M83749; AAA37519.1; -;  
DR EMBL; M86182; AAA37503.1; -;  
DR PIR; B40035; B40035.  
DR PIR; A41984; A41984.  
DR MGD; MGI:88314; Ccnd2.  
DR InterPro; IPR004366; Cyclin.  
DR InterPro; IPR004367; Cyclin\_Cterm.  
DR Pfam; PF00134; cyclin; 1.  
DR Pfam; PF02984; cyclin\_C; 1.  
DR SMART; SM00385; CYCLIN; 1.  
DR PROSITE; PS00292; CYCLINS; 1.  
DR Cyclin; Cell cycle; Cell division; Multigene family.  
KW SEQUENCE 289 AA; 32897 MW; 58F322771DD1DA3D CRC64;  
SQ  
  
Query Match 61.5%; Score 32; DB 1; Length 289;  
Best Local Similarity 60.0%; Pred. No. 29;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
  
QY 1 EEVVPXGMSY 10  
DQ ||||| |  
Db 73 EEVFPPLAMNY 82  
  
RESULT 21  
CGD1\_BRARE  
ID CGD1\_BRARE STANDARD; PRT; 291 AA.  
AC Q90459;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE G1/S-specific cyclin D1.  
GN CYCD1.  
OS Brachydanio rerio (Zebrafish) (Danio rerio).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
OC Cyprinidae; Danio.  
OX NCBI\_TaxID=7955;  
RN [1]

RP SEQUENCE FROM N.A.  
 RX MEDLINE=96138542; PubMed=8547308;  
 RA Yarden A., Salomon D., Geiger B.;  
 RT "Zebrafish cyclin D1 is differentially expressed during early  
 RT embryogenesis.";   
 RL Biochim. Biophys. Acta 1264:257-260(1995).  
 CC -1- FUNCTION: ESSENTIAL FOR THE CONTROL OF THE CELL CYCLE AT THE G1/S  
 CC (START) TRANSITION.  
 CC -1- SUBUNIT: INTERACTS WITH THE CDK4 AND CDK6 PROTEIN KINASES TO FORM  
 CC A SERINE/THREONINE KINASE HOLOENZYME COMPLEX. THE CYCLIN SUBUNIT  
 CC IMPARTS SUBSTRATE SPECIFICITY TO THE COMPLEX.  
 CC -1- SIMILARITY: BELONGS TO THE CYCLIN FAMILY. CYCLIN D SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X87581; CAA60885.1; -;  
 DR ZFIN; ZDB-GENE-980526-176; cycdl.  
 DR InterPro; IPR004366; Cyclin.  
 DR Pfam; PF00134; cyclin; 1.  
 DR SMART; SM00385; cyclin.C; 1.  
 DR PROSITE; PS00292; CYCLIN; 1.  
 DR CYCLIN; Cell cycle; Cell division.  
 KW CYCLIN; Cell cycle; Cell division.  
 SQ SEQUENCE 291 AA; 33067 MW; FA5274CB1B46D5EF CRC64;  
 -----  
 Query Match 61.5%; Score 32; DB 1; Length 291;  
 Best Local Similarity 60.0%; Pred. No. 29;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 EEVVPXGMSY 10  
 DB ||||| |:  
 75 EEVFPPLAMNY 84  
 -----  
 RESULT 22  
 CGDL\_XENLA  
 ID CGDL\_XENLA STANDARD; PRT; 291 AA.  
 AC P50755;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE G1/S-specific cyclin D1.  
 GN CCND1.  
 OS Xenopus laevis (African clawed frog).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;  
 CC Xenopodinae; Xenopus.  
 CC NCBI\_TaxID=8355;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Cockerill M.J., Hunt T.;  
 RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: ESSENTIAL FOR THE CONTROL OF THE CELL CYCLE AT THE G1/S  
 CC (START) TRANSITION.  
 CC -1- SUBUNIT: INTERACTS WITH THE CDK4 AND CDK6 PROTEIN KINASES TO FORM  
 CC A SERINE/THREONINE KINASE HOLOENZYME COMPLEX. THE CYCLIN SUBUNIT  
 CC IMPARTS SUBSTRATE SPECIFICITY TO THE COMPLEX.  
 CC -1- SIMILARITY: BELONGS TO THE CYCLIN FAMILY. CYCLIN D SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

CC EMBL; X89475; CAA61664.1; -;  
 DR InterPro; IPR004366; Cyclin.  
 DR Pfam; PF00134; cyclin; 1.  
 DR SMART; SM00385; cyclin.C; 1.  
 DR PROSITE; PS00292; CYCLIN; 1.  
 DR CYCLIN; Cell cycle; Cell division; Multigene family.  
 KW CYCLIN; Cell cycle; Cell division; Multigene family.  
 SQ SEQUENCE 291 AA; 32953 MW; A4747C5BD1679087 CRC64;  
 -----  
 Query Match 61.5%; Score 32; DB 1; Length 291;  
 Best Local Similarity 60.0%; Pred. No. 29;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 EEVVPXGMSY 10  
 DB ||||| |:  
 73 EEVFPPLAMNY 82  
 -----  
 RESULT 23  
 CGD2\_CHICK  
 ID CGD2\_CHICK STANDARD; PRT; 291 AA.  
 AC P49706;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE G1/S-specific cyclin D2.  
 GN CCND2.  
 OS Gallus gallus (Chicken).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
 CC Gallus.  
 CC NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=96144302; PubMed=8566807;  
 RA Li H., Grenet J., Kidd V.J.;  
 RT "Structure and gene expression of avian cyclin D2";  
 RL Gene 167:341-342(1995).  
 CC -1- FUNCTION: ESSENTIAL FOR THE CONTROL OF THE CELL CYCLE AT THE G1/S  
 CC (START) TRANSITION.  
 CC -1- SUBUNIT: INTERACTS WITH THE CDK4 AND CDK6 PROTEIN KINASES TO FORM  
 CC A SERINE/THREONINE KINASE HOLOENZYME COMPLEX. THE CYCLIN SUBUNIT  
 CC IMPARTS SUBSTRATE SPECIFICITY TO THE COMPLEX.  
 CC -1- SIMILARITY: BELONGS TO THE CYCLIN FAMILY. CYCLIN D SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; U28980; AAA96955.1; -;  
 DR InterPro; IPR004366; Cyclin.  
 DR Pfam; PF00134; cyclin; 1.  
 DR SMART; SM00385; cyclin.C; 1.  
 DR PROSITE; PS00292; CYCLIN; 1.  
 DR CYCLIN; Cell cycle; Cell division; Multigene family.  
 KW CYCLIN; Cell cycle; Cell division; Multigene family.  
 SQ SEQUENCE 291 AA; 33163 MW; 43A7E646AEAF3109 CRC64;  
 -----  
 Query Match 61.5%; Score 32; DB 1; Length 291;  
 Best Local Similarity 60.0%; Pred. No. 29;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 EEVVPXGMSY 10  
 DB ||||| |:  
 74 EEVFPPLAMNY 83

## RESULT 24

CGD2\_XENLA STANDARD; PRT; 291 AA.  
ID CGD2\_XENLA  
AC P53782;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE GL/S-specific cyclin D2.  
GN CCND2.  
OS Xenopus laevis (African clawed frog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;  
OC Xenopodinae; Xenopus.  
OX NCBI\_TaxID=8355;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Cockerill M.J., Hunt T.;  
RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.  
[2]  
RN  
RP SEQUENCE FROM N.A.  
RX MEDLINE-97380591; PubMed-9237366;  
RA Taieb F., Jessus C.;  
RT "Xenopus cyclin D2: cloning and expression in oocytes and during  
early development.";  
RL Biol. Cell 88:99-111(1996).  
CC -!- FUNCTION: ESSENTIAL FOR THE CONTROL OF THE CELL CYCLE AT THE G1/S  
(START) TRANSITION.  
CC -!- SUBUNIT: INTERACTS WITH THE CDK4 AND CDK6 PROTEIN KINASES TO FORM  
A SERINE/THREONINE KINASE HOLOENZYME COMPLEX. THE CYCLIN SUBUNIT  
IMPARTS SUBSTRATE SPECIFICITY TO THE COMPLEX.  
CC -!- SIMILARITY: BELONGS TO THE CYCLIN FAMILY. CYCLIN D SUBFAMILY.

-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL; X89476; CAA61665.1; -;  
DR EMBL; X83503; CAA58493.1; -;  
DR InterPro; IPR004366; Cyclin.  
DR InterPro; IPR004367; Cyclin\_Cterm.  
DR Pfam; PF00134; cyclin; 1.  
DR Pfam; PF02984; cyclin\_C; 1.  
DR SMART; SM00385; CYCLIN; 1.  
DR PROSITE; PS00292; CYCLINS; 1.  
KW Cyclin; Cell cycle; Cell division; Multigene family.  
SQ SEQUENCE 291 AA; 32959 MW; 9A290F04F1531E89 CRC64;

Query Match 61.5%; Score 32; DB 1; Length 291;  
Best Local Similarity 60.0%; Pred. No. 29;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 EEVVPXGMSY 10  
||| | | |  
DB 74 EEVFPAMNY 83

## RESULT 25

CGD1\_CHICK STANDARD; PRT; 292 AA.  
ID CGD1\_CHICK  
AC P5169;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE GL/S-specific cyclin D1.  
GN CCND1.  
OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;

OC Gallus.  
OX NCBI\_TaxID=9031;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Li H., Lahti J.M., Kidd V.J.;  
RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: ESSENTIAL FOR THE CONTROL OF THE CELL CYCLE AT THE G1/S  
(START) TRANSITION. INTERACTS WITH THE CDK4 AND CDK6 PROTEIN  
KINASES (BY SIMILARITY).  
CC -!- SIMILARITY: BELONGS TO THE CYCLIN FAMILY. CYCLIN D SUBFAMILY.

-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL; U40844; AA83271.1; -;  
DR InterPro; IPR004366; Cyclin.  
DR InterPro; IPR004367; Cyclin\_Cterm.  
DR Pfam; PF00134; cyclin; 1.  
DR Pfam; PF02984; cyclin\_C; 1.  
DR SMART; SM00385; CYCLIN; 1.  
DR PROSITE; PS00292; CYCLINS; 1.  
KW Cyclin; Cell cycle; Cell division; Multigene family.  
SQ SEQUENCE 292 AA; 33263 MW; 7B543029DB45A67D CRC64;

Query Match 61.5%; Score 32; DB 1; Length 292;  
Best Local Similarity 60.0%; Pred. No. 30;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 EEVVPXGMSY 10  
||| | | |  
DB 75 EEVFPAMNY 84

Search completed: June 10, 2003, 13:40:17  
Job time : 5.5 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2003, 13:30:25 ; Search time 25.7857 Seconds  
(without alignments)  
87.898 Million cell updates/sec

Title: US-09-909-164-6  
Perfect score: 52  
Sequence: 1 EHVVPXGMSYS 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 75 summaries

Database : SPTREMBL\_21.\*

- 1: sp\_archaea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mhc.\*
- 8: sp\_organelle.\*
- 9: sp\_phage.\*
- 10: sp\_plant.\*
- 11: sp\_rodent.\*
- 12: sp\_virus.\*
- 13: sp\_vertebrate.\*
- 14: sp\_unclassified.\*
- 15: sp\_virus.\*
- 16: sp\_bacteria.\*
- 17: sp\_archaea.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	38	73.1	387	16	Q98FX1	Q98fx1 rhizobium 1
2	38	73.1	1063	16	Q8RG86	Q8rg86 fusobacteri
3	38	73.1	3472	1	O74056	O74056 cenarchaeum
4	37	71.2	840	3	Q9URY8	Q9ury8 schizosacch
5	36	69.2	471	11	Q8R126	Q8r126 mus musculus
6	36	69.2	484	11	Q8VD18	Q8vd18 mus musculus
7	35	67.3	225	10	Q40129	Q40129 lycopersico
8	35	67.3	425	5	Q9XVK4	Q9xvk4 caenorhabdi
9	35	67.3	556	4	O43733	O43733 homo sapien
10	35	67.3	583	5	Q9BHA5	Q9bha5 plasmodium
11	35	67.3	583	5	Q9BH83	Q9bh83 plasmodium
12	35	67.3	670	11	Q01487	Q01487 rattus norv
13	35	67.3	749	16	Q9PDM6	Q9pdm6 xylella fas
14	35	67.3	1902	4	Q14122	Q14122 homo sapien
15	34	65.4	156	3	Q12479	Q12479 saccharomyc
16	34	65.4	219	17	Q971S2	Q971s2 sulfolobus

17	34	65.4	252	17	O28342	O28342 archaeoglob
18	34	65.4	290	4	Q96MU1	Q96mul homo sapien
19	34	65.4	387	16	Q92MD6	Q92md6 rhizobium m
20	34	65.4	541	16	Q98BP5	Q98bp5 rhizobium 1
21	34	65.4	544	16	Q9PQD2	Q9pqd2 ureaplasma
22	34	65.4	842	3	Q9URR4	Q9urr4 penicillium
23	34	65.4	1049	16	Q8XT05	Q8xt05 ralstonia s
24	34	65.4	1499	4	Q96914	Q96914 homo sapien
25	33	63.5	143	17	Q8TX62	Q8tx62 methanopyru
26	33	63.5	162	11	Q9CXQ4	Q9cxq4 mus musculus
27	33	63.5	165	17	O28330	O28330 archaeoglob
28	33	63.5	193	2	Q8VUA8	Q8vuua lactococcus
29	33	63.5	209	16	Q8RE56	Q8re56 fusobacteri
30	33	63.5	284	16	P74187	P74187 synechocyst
31	33	63.5	298	10	Q9M3C0	Q9m3c0 arabidopsis
32	33	63.5	326	12	Q9Q9Q9	Q9q9q9 soil-borne
33	33	63.5	326	12	Q9Q9Q5	Q9q9q5 soil-borne
34	33	63.5	326	12	Q9QCE7	Q9qce7 soil-borne
35	33	63.5	326	12	Q9DJG4	Q9djg4 soil-borne
36	33	63.5	326	12	Q91DN1	Q91dn1 soil-borne
37	33	63.5	326	12	Q9Q9Q7	Q9q9q7 soil-borne
38	33	63.5	327	12	Q06360	Q06360 soil-borne
39	33	63.5	332	10	Q9FNL4	Q9fnl4 arabidopsis
40	33	63.5	368	16	Q9X0U3	Q9x0u3 thermotoga
41	33	63.5	393	5	Q9V914	Q9v914 drosophila
42	33	63.5	479	4	Q96CS0	Q96cs0 homo sapien
43	33	63.5	548	11	Q9D2X9	Q9d2x9 mus musculus
44	33	63.5	584	16	Q8F8K6	Q8f8k6 thermoanaer
45	33	63.5	648	4	Q96MB2	Q96mb2 homo sapien
46	33	63.5	653	16	Q9KVE3	Q9kve3 vibrio chol
47	33	63.5	676	5	Q9VA55	Q9va55 drosophila
48	33	63.5	678	12	Q8T8Z7	Q8t8z7 drosophila
49	33	63.5	746	3	Q9URR3	Q9urr3 cercopithec
50	33	63.5	791	4	Q9H2K5	Q9h2k5 homo sapien
51	33	63.5	793	4	Q9H2K6	Q9h2k6 homo sapien
52	33	63.5	844	11	Q922D4	Q922d4 mus musculus
53	33	63.5	1028	16	Q8YJ11	Q8yjl1 bruceella me
54	33	63.5	1152	16	Q9Q9C9	Q9qc95 mycobacteri
55	33	63.5	1305	5	Q9V7C7	Q9v7c7 drosophila
56	33	63.5	1394	4	Q8TD95	Q8td95 homo sapien
57	33	63.5	1442	17	Q96YH5	Q96yh5 sulfolobus
58	33	63.5	1548	10	O65531	O65531 arabidopsis
59	33	63.5	1713	11	O88349	O88349 mus musculus
60	33	63.5	84	16	Q97DE7	Q97de7 clostridium
61	32	61.5	103	11	Q9D0H9	Q9d0h9 mus musculus
62	32	61.5	108	1	Q9UX33	Q9ux33 sulfolobus
63	32	61.5	108	1	Q9UX33	Q9ux33 sulfolobus
64	32	61.5	133	13	P79919	P79919 xenopus lae
65	32	61.5	156	11	Q9D8L9	Q9d8l9 mus musculus
66	32	61.5	174	10	Q9M3T4	Q9m3t4 betula verr
67	32	61.5	190	13	O57481	O57481 stizostedio
68	32	61.5	191	11	Q99NB4	Q99nb4 rattus norv
69	32	61.5	200	17	Q97CD0	Q97cd0 thermoplas
70	32	61.5	207	2	Q47284	Q47284 escherichia
71	32	61.5	234	2	O32330	O32330 clostridium
72	32	61.5	236	10	Q9SXF1	Q9sxf1 arabidopsis
73	32	61.5	240	11	Q9DB09	Q9db09 mus musculus
74	32	61.5	243	12	Q91EW1	Q91ew1 cydia pomon
75	32	61.5	279	17	Q9Y8Z4	Q9y8z4 aeropyrum p

## ALIGNMENTS

RESULT 1  
Q98FX1 PRELIMINARY; PRT; 387 AA.  
ID Q98FX1  
AC Q98FX1;  
DT 01-OCT-2001 (TrEMBLrel. 18, Created)  
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)  
DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)  
DE Hippurate hydrolase.  
GN MLR3583.

```

OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-NAFF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsumoto A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti.";
RL DNA Res 7:331-338(2000).
DR EMBL: AF003002; BAB50445.1; -.
DR InterPro: IPR002933; Peptidase_M20.
DR Pfam: PF01546; Peptidase_M20; 1.
DR Hydrolase; Complete proteome.
SQ SEQUENCE 387 AA; 41180 MW; 131BFF8E64306829 CRC64;

Query Match 73.1%; Score 38; DB 16; Length 387;
Best Local Similarity 60.0%; Pred. No. 9.8;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 EVVVPXGMSY 10
Db 367 DEAIPIHGMYS 376

RESULT 2
Q8RG86
ID Q8RG86 PRELIMINARY; PRT; 1063 AA.
AC Q8RG86;
DT 01-JUN-2002 (TReMBLrel. 21, Created)
DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Carbamoyl-phosphate synthase large chain (EC 6.3.5.5).
GN FN0422.
OS Fusobacterium nucleatum (subsp. nucleatum).
OC Bacteria; Fusobacteria; Fusobacterium.
OX NCBI_TaxID=76856;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 25586;
RX MEDLINE=21886394; PubMed=11889109;
RA Bhattacharya A., Bartman A., Gardner W., Grechkin G., Zhu L.,
RA Vasieva O., Chu L., Kogan Y., Chaga O., Goltsman E., Bernal A.,
RA Larsen N., D'Souza M., Walunas T., Pusch G., Haselkorn R.,
RA Fontstein M., Kypides N., Overbeek R.;
RT "Genome sequence and analysis of the oral bacterium Fusobacterium
RT nucleatum strain ATCC 25586.";
RL J. Bacteriol. 184:2005-2018(2002).
DR EMBL: AE010554; AAL94625.1; -.
DR Ligase; Complete proteome.
SQ SEQUENCE 1063 AA; 118008 MW; 39700E10B7CCE411 CRC64;

Query Match 73.1%; Score 38; DB 16; Length 1063;
Best Local Similarity 60.0%; Pred. No. 29;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 EVVVPXGMSYS 11
Db 195 EIVPNGINYS 204

RESULT 3
O74056
ID O74056 PRELIMINARY; PRT; 3472 AA.
AC O74056;
DT 01-NOV-1998 (TReMBLrel. 08, Created)

```

```

DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Hypothetical 367.1 kDa protein.
OS Cenarchaeum symbiosum.
OC Archaea; Crenarchaeota; Cenarchaeum.
OX NCBI_TaxID=46770;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=B;
RX MEDLINE=98422450; PubMed=9748430;
RA Schleper C., DeLong E.F., Preston C.M., Feldman R.A., Wu K.Y.,
RA Swanson R.V.;
RT "Genomic analysis reveals chromosomal variation in natural populations
RT of the uncultured psychrophilic archaeon Cenarchaeum symbiosum.";
RL J. Bacteriol. 180:5003-5009(1998).
CC -1- SIMILARITY: CONTAINS 5 WD REPEATS (TRP-ASP DOMAINS).
DR EMBL: AF083072; AAC62699.1; -.
DR InterPro: IPR000515; BPD_transp.
DR Pfam: PF00400; WD40; 4.
DR SMART: SM00320; WD40; 2.
DR PROSITE: PS00402; BPD_TRANSF_INN_MEMBER; UNKNOWN_1.
KW Hypothetical protein; Repeat; WD repeat.
SQ SEQUENCE 3472 AA; 367058 MW; 37F80707030F9355 CRC64;

Query Match 73.1%; Score 38; DB 1; Length 3472;
Best Local Similarity 54.5%; Pred. No. 1e+02;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 EVVVPXGMSYS 11
Db 2294 EDVIPRGISFS 2304

RESULT 4
Q9URY8
ID Q9URY8 PRELIMINARY; PRT; 840 AA.
AC Q9URY8;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Probable sulfate permease.
GN SPAC869.05C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972H-;
RA Hunt C., Aves S., McDougall R.C., Rajandream M.A., Barrell B.G.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL132779; CAB60015.1; -.
DR InterPro: IPR002645; STAS.
DR InterPro: IPR001902; Sulfate_transp.
DR Pfam: PF01740; STAS; 1.
DR Pfam: PF00916; Sulfate_transp; 1.
DR TIGRFAMs: TIGR00815; sulp; 1.
DR SEQUENCE 840 AA; 93517 MW; ED4833E162B69077 CRC64;

Query Match 71.2%; Score 37; DB 3; Length 840;
Best Local Similarity 77.8%; Pred. No. 37;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMSYS 11
Db 135 VVPQGMYSY 143

RESULT 5
Q8R126
ID Q8R126 PRELIMINARY; PRT; 471 AA.

```



AC Q8RL26;  
 DT 01-JUN-2002 (TRENBLrel. 21, Created)  
 DT 01-JUN-2002 (TRENBLrel. 21, Last sequence update)  
 DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)  
 DE Hypothetical 54.5 kDa protein (Fragment).  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=LIVER;  
 RA Strausberg R.;  
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC025810; AAH25810.1; -;  
 KW Hypothetical protein.  
 FT NON\_TER 1  
 SQ SEQUENCE 471 AA; 54506 MW; E0DA685C374A9760 CRC64;

Query Match 69.2%; Score 36; DB 11; Length 471;  
 Best Local Similarity 60.0%; Pred. No. 32;  
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 EVVPXGMSYS 11  
 DB 226 EVIPAGASYN 235  
 II: I I I I

RESULT 6  
 Q8VD18 PRELIMINARY; PRT; 484 AA.  
 ID Q8VD18;  
 AC Q8VD18;  
 DT 01-MAR-2002 (TRENBLrel. 20, Created)  
 DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)  
 DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)  
 DE Similar to glioma tumor suppressor candidate region gene 2.  
 GN AW536441.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=SALIVARY GLAND;  
 RA Strausberg R.;  
 RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC017637; AAH17637.1; -;  
 DR MGD; MGI:2138595; AW536441.  
 SQ SEQUENCE 484 AA; 55835 MW; BBB45F3B4BE02A36 CRC64;

Query Match 69.2%; Score 36; DB 11; Length 484;  
 Best Local Similarity 60.0%; Pred. No. 33;  
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 EVVPXGMSYS 11  
 DB 239 EVIPAGASYN 248  
 II: I I I I

RESULT 7  
 Q40129 PRELIMINARY; PRT; 225 AA.  
 ID Q40129;  
 AC Q40129;  
 DT 01-NOV-1996 (TRENBLrel. 01, Created)  
 DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)  
 DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)  
 DE Hypothetical 25.2 kDa protein precursor.  
 OS Lycopersicon esculentum (Tomato).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.  
 OX NCBI\_TaxID=4081;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RC STRAIN=VF36; TISSUE=PISTIL;  
 RX MEDLINE=95375233; PubMed=7647301;  
 RA Milligan S.B.; Gasser C.S.;  
 RT "Nature and regulation of pistil-expressed genes in tomato.";  
 RL Plant Mol. Biol. 28:691-711(1995).  
 DR EMBL; U20592; AAA80497.1; -;  
 DR InterPro; IPR002160; Kunitz\_legume.  
 DR Pfam; PF00197; Kunitz\_legume; 1.  
 DR ProDom; PD000891; Kunitz\_legume; 1.  
 DR ProDom; SM00452; STI; 1.  
 DR DR PROSITE; PS00283; SOYBEAN\_KUNITZ; UNKNOWN\_1.  
 KW Hypothetical protein; Signal.  
 FT CHAIN 1 20 POTENTIAL.  
 FT CHAIN 21 225 UNKNOWN.  
 SQ SEQUENCE 225 AA; 25188 MW; 1074C261D20CFDAD CRC64;

Query Match 67.3%; Score 35; DB 10; Length 225;  
 Best Local Similarity 54.5%; Pred. No. 24;  
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMSYS 11  
 DB 32 DEVVPNGKTYA 42  
 : I I I I I I I I

RESULT 8  
 Q9XVK4 PRELIMINARY; PRT; 425 AA.  
 ID Q9XVK4;  
 AC Q9XVK4;  
 DT 01-NOV-1999 (TRENBLrel. 12, Created)  
 DT 01-NOV-1999 (TRENBLrel. 12, Last sequence update)  
 DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)  
 DE R10D12.10 protein.  
 GN R10D12.10.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Percy C.M.;  
 RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RX MEDLINE=99069613; PubMed=9851916;  
 RA none;  
 RT "Genome sequence of the nematode C.elegans: A platform for  
 investigating biology.";  
 RL Science 282:2012-2018(1998).  
 DR EMBL; Z81109; CAB03241.1; -;  
 DR InterPro; IPR000719; Euk\_pkinase.  
 DR Pfam; PF00069; pkinase; 1.  
 DR ProDom; PD000001; Euk\_pkinase; 1.  
 DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
 KW ATP-binding; Transferase.  
 SQ SEQUENCE 425 AA; 49410 MW; 5D96E29B08C8E9D6 CRC64;

Query Match 67.3%; Score 35; DB 5; Length 425;  
 Best Local Similarity 50.0%; Pred. No. 48;  
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10  
 DB 335 EQIVPGGLQY 344  
 I: I I I I I

RESULT 9  
 O43733 PRELIMINARY; PRT; 556 AA.  
 ID O43733;  
 AC O43733;  
 DT 01-JUN-1998 (TRENBLrel. 06, Created)  
 DT 01-JUN-1998 (TRENBLrel. 06, Last sequence update)

DT 01-JUN-1998 (Tremblrel. 06, Last annotation update)  
 DE DNA binding protein (Fragment).  
 GN D3451B15.2.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Tubby B.;  
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; Z98050; CAB10847.1; -.  
 FT NON\_TER 1  
 SQ SEQUENCE 556 AA; 59059 MW; ECB00E403FB2528 CRC64;

Query Match 67.3%; Score 35; DB 4; Length 556;  
 Best Local Similarity 66.7%; Pred. No. 64;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMSYS 11  
 ||| |::||  
 Db 244 VVPAGLTYS 252

RESULT 10  
 Q9BHA5 PRELIMINARY; PRT; 583 AA.  
 AC Q9BHA5;  
 DT 01-JUN-2001 (Tremblrel. 17, Created)  
 DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)  
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)  
 DE Choline transporter.  
 GN SCT1.  
 OS Plasmodium falciparum.  
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
 OX NCBI\_TaxID=5833;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Ben Mamoun C., Gluzman I.Y., Goldberg D.E.;  
 RT "Plasmodium falciparum choline transporter (pfSCT1) gene.";  
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY007372; AAK14815.1; -.  
 DR EMBL; AY007375; AAG17947.1; -.  
 DR InterPro; IPR002123; Acyltransferase.  
 DR Pfam; PF01553; Acyltransferase; 1.  
 SQ SEQUENCE 583 AA; 66917 MW; 2B2BFAE3E395E049 CRC64;

Query Match 67.3%; Score 35; DB 5; Length 583;  
 Best Local Similarity 55.6%; Pred. No. 67;  
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMSYS 11  
 ||| |::||  
 Db 227 IIPVGLSYS 235

RESULT 11  
 Q9BH83 PRELIMINARY; PRT; 583 AA.  
 AC Q9BH83;  
 DT 01-JUN-2001 (Tremblrel. 17, Created)  
 DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)  
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)  
 DE Choline transporter.  
 GN SCT1.  
 OS Plasmodium falciparum.  
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
 OX NCBI\_TaxID=5833;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Ben Mamoun C., Gluzman I.Y., Goldberg D.E.;  
 RT "Plasmodium falciparum choline transporter (pfSCT1) gene.";  
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AY007374; AAK14818.1; -.  
 DR EMBL; AY007373; AAK14817.1; -.  
 DR InterPro; IPR002123; Acyltransferase.  
 DR Pfam; PF01553; Acyltransferase; 1.  
 SQ SEQUENCE 583 AA; 66918 MW; 4DF83D7530527474 CRC64;

Query Match 67.3%; Score 35; DB 5; Length 583;  
 Best Local Similarity 55.6%; Pred. No. 67;  
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMSYS 11  
 ||| |::||  
 Db 227 IIPVGLSYS 235

RESULT 12  
 Q01487 PRELIMINARY; PRT; 670 AA.  
 ID Q01487;  
 AC Q01487;  
 DT 01-NOV-1996 (Tremblrel. 01, Created)  
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)  
 DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)  
 DE DNA-binding protein AT-BP2 (ALPHA1-antitrypsin promoter binding protein 2) (Fragment).  
 DE protein 2) (Fragment).  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-SPRAGUE-DAWLEY; TISSUE-THYROID;  
 RX MEDLINE=91187610; PubMed=1901405;  
 RA Mitchelmore C., Traboni C., Cortese R.;  
 RT "Isolation of two cDNAs encoding zinc finger proteins which bind to the alpha 1-antitrypsin promoter and to the major histocompatibility complex class I enhancer.";  
 RT the alpha 1-antitrypsin promoter and to the major histocompatibility complex class I enhancer.";  
 RL Nucleic Acids Res. 19:141-147(1991).  
 CC -1- FUNCTION: BINDS TO THE ALPHA1-ANTITRYPSIN PROMOTER, TO THE KAPPA IMMUNOGLOBULIN GENE ENHANCER, AND TO THE MAJOR HISTOCOMPATIBILITY COMPLEX CLASS I ENHANCER: IT PLAYS A ROLE AS TRANSCRIPTIONAL REGULATOR. CONCERNING THE ALPHA1-ANTITRYPSIN EXPRESSION IT MAY ACT AS A NEGATIVE REGULATOR BY INTERFERING AT THE ALPHA1-ANTITRYPSIN PROMOTOR WITH ANOTHER NUCLEAR PROTEIN, CALLED LF-B.  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR.  
 CC -1- MISCELLANEOUS: ACIDIC AMINO ACID REGION MAY BE INVOLVED IN TRANSCRIPTIONAL ACTIVATION.  
 CC -1- SIMILARITY: HIGHLY SIMILAR TO AT-BP1 AND TO THE PRDII-BF1 FACTOR.  
 DR EMBL; X54250; CAA38151.1; -.  
 DR HSPSP; P15822; IBB0.  
 DR InterPro; IPR000822; Znf\_C2H2.  
 DR Pfam; PF00096; zf-C2H2; 2.  
 DR SMART; SM00355; Znf\_C2H2; 2.  
 DR PROSITE; PS00028; ZINC\_FINGER\_C2H2\_1; 2.  
 DR PROSITE; PS0157; ZINC\_FINGER\_C2H2\_2; 2.  
 KW Zinc-finger; Nuclear protein; DNA-binding; Transcription regulation;  
 KW Metal-binding; Multigene family.  
 FT NON\_TER 1  
 FT DOMAIN 54 104 ZINC-FINGERS.  
 FT DOMAIN 140 160 ACIDIC.  
 FT ZN\_FING 54 74 C(2)H(2) CLASS.  
 FT ZN\_FING 82 104 C(2)H(2) CLASS.  
 SQ SEQUENCE 670 AA; 71233 MW; CDD2324152590C17 CRC64;

Query Match 67.3%; Score 35; DB 11; Length 670;  
 Best Local Similarity 66.7%; Pred. No. 78;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMSYS 11  
 ||| |::||  
 Db 376 VVPAGLTYS 384

RESULT 13

OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.

AC MEDLINE=90205817; PubMed=2108316;  
DT Baldwin A.S., Leclair K.P., Singh H., Sharp P.A.;  
DT "A large protein containing zinc finger domains binds to related  
DT sequence elements in the enhancers of the class I major  
DE histocompatibility complex and kappa immunoglobulin genes.";  
GN Mol. Cell. Biol. 10:1406-1414 (1990).  
OS MBL; M32019; AAA17534.1;  
OC HSP; P15822; 1BBO.  
OX InterPro: IPR000822; Znf\_C2H2.  
RN Pfam: PF00096; Znf\_C2H2; 2.  
RP SMART: SM00355; Znf\_C2H2; 2.  
AC PROSITE: PS00028; ZINC\_FINGER\_C2H2\_1; 2.  
DT PROSITE: PS0157; ZINC\_FINGER\_C2H2\_2; 2.  
DE DNA-binding; Metal-binding; Zinc-finger.  
GN NON\_TER 1  
OS SEQUENCE 1902 AA; 207457 MW; 35BBF0D961E6ED20 CRC64;  
OC Query Match 67.3%; Score 35; DB 4; Length 1902;  
OX Best Local Similarity 66.7%; Pred. No. 2.4e+02;  
RN Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMSYS 11  
Db 1590 VVPAGLTVS 1598  
PRT; 156 AA.

RESULT 15  
Q12479 PRELIMINARY;  
ID Q12479 PRELIMINARY;  
AC Q12479  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE YOR013W.  
GN YOR013W.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
OX NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA De haan M., Grivell L.A., Maarse A.C.;  
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA MIPS;  
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RA STRAIN-FY1679;  
RL De haan M., Maarse A.C., Grivell L.A.;  
RN Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A.  
RA STRAIN-FY1679;  
RL MEDLINE=94019318; PubMed=8413243;  
RN Dumont M.E., Schlichter J.B., Cardillo T.S., Hayes M.K., Bethlenny G.,  
RA Sherman F.;  
RT "CYC2 encodes a factor involved in mitochondrial import of yeast  
RT cytochrome c.";  
RL Mol. Cell. Biol. 13:6442-6451 (1993).  
RN [5]  
RP SEQUENCE FROM N.A.  
RC STRAIN-FY1679;  
RX MEDLINE=94169519; PubMed=7764548;  
RA Lee Y.S., Shimizu J., Yoda K., Yamasaki M.;  
RT "Molecular cloning of a gene, DHS1, which complements a drug-  
RT hypersensitive mutation of the yeast Saccharomyces cerevisiae.";  
RL Biosci. Biotechnol. Biochem. 58:391-395 (1994).

Q9PDM6 PRELIMINARY; PRT; 749-AA.  
AC Q9PDM6;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE Topoisomerase IV subunit.  
GN XFL353.  
OS Xylella fastidiosa.  
OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;  
OX NCBI\_TaxID=2371;  
RN [1]  
RP SEQUENCE FROM N.A.

AC MEDLINE=945C5717; PubMed=10910347;  
DT Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,  
DT Alvaranga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,  
DT Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,  
DT Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carrer H.,  
DT Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,  
DT Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorry H.,  
DT Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,  
DT Fraga J.S., Franca S.C., Franco M.C., Frohne M., Furlan L.R.,  
DT Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,  
DT Ho P.L., Hohelsel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,  
DT Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,  
DT Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,  
DT Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,  
DT Marques M.V., Miracca E.C., Miyaki C.F., Monteiro-Vitorello C.B.,  
DT Menck C.F.M., Miracca E.C., Miyaki C.F., Monteiro-Vitorello C.B.,  
DT Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,  
DT Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,  
DT de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,  
DT Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,  
DT Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,  
DT de Rosa V.B. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,  
DT da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,  
DT da Silva J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,  
DT de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tshunoko M.H.,  
DT Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,  
DT Zago M.A., Zatz M., Meidanis J., Setubal J.C.;  
RT "The genome sequence of the plant pathogen Xylella fastidiosa.";  
RL Nature 406:151-159 (2000).  
OC EMBL; AE003967; AAF84162.1;  
OX HSP; P09097; 1AB4.  
RN InterPro: IPR002205; DNA\_topoisomIV.  
RP Pfam: PF00521; DNA\_topoisomIV; 1.  
DT ProDom: PD000742; DNA\_topoisomIV; 1.  
DE SMART: SM00434; TOP4C; 1.  
GN TIGRFAMs: TIGR01062; parC\_Gneg; 1.  
OS Complete proteome.  
OC SEQUENCE 749 AA; 83344 MW; 4DCD10F40EE0257 CRC64;

Query Match 67.3%; Score 35; DB 16; Length 749;  
Best Local Similarity 77.8%; Pred. No. 88;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 EVVPXGMSYS 10  
Db 526 EVDPSGMSYS 534  
PRT; 1902 AA.

RESULT 14  
Q14122 PRELIMINARY;  
ID Q14122  
AC Q14122;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
DE DNA-binding protein (Mdp-1) (Fragment).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

DR EMBL: 274920; CAA99201.1; -
DR EMBL: X87331; CAA60762.1; -
DR SGD: S0005539; YOR013W.
SQ SEQUENCE 156 AA; 17881 MW; 380442B74C272B41 CRC64;

Query Match
Best Local Similarity 65.4%; Score 34; DB 3; Length 156;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 EVVPXGMSY 10
Db 50 EVPLGMDY 58

RESULT 16
Q971S2 PRELIMINARY; PRT; 219 AA.
AC Q971S2;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DE Putative ribose 5-phosphate isomerase.
GN SRI302.
OS Sulfolobus tokodaii.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.
OC NCBI_TaxID=111955;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JCM 10545 / 7;
RX PubMed=11572479;
RA Kawarabayashi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,
RA Sekine M., Baba S.-I., Ankaï A., Kosugi H., Hosoyama A., Fukui S.,
RA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,
RA Yoshizawa T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,
RA Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A.,
RA Oshima T., Kikuchi H.;
RT "Complete genome sequence of an aerobic thermoacidophilic
RT Crenarchaeon, Sulfolobus tokodaii strain7.";
RL DNA Res. 8:123-140(2001).
DR EMBL: AP000985; BAB66348.1; -
DR InterPro: IPR004788; RPIA.
DR ProDom: PD005813; RPIA; 1.
KW Isomerase; Hypothetical protein; Complete proteome.
SQ SEQUENCE 219 AA; 24541 MW; A4E9A3F2C4006D90 CRC64;

Query Match
Best Local Similarity 65.4%; Score 34; DB 17; Length 219;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 EVVPXGMSY 10
Db 131 EVVPGVAY 139

RESULT 17
O28342 PRELIMINARY; PRT; 252 AA.
ID O28342
AC O28342;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Cell division inhibitor (MIND-2).
GN AF1937.
OS Archaeoglobus fulgidus.
OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
OC Archaeoglobaceae; Archaeoglobus.
OC NCBI_TaxID=2234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=98049343; PubMed=9389475;
RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,

Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
Richardson D.L., Kerlavage A.R., Graham D.E., Kyrpides N.C.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
RA Cotton M.D., Spriggs T., Artiach P., Kaine B.P., Sykes S.M.,
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RT "The complete genome sequence of the hyperthermophilic, sulphate-
RT reducing archaeon Archaeoglobus fulgidus.";
RL Nature 390:364-370(1997).
DR EMBL: AE000970; AAB89318.1; -
DR TIGR: AF1937; -
DR InterPro: IPR000707; ATPase_Para.
DR Pfam: PF00991; Para; 1.
KW Hypothetical protein; Cell division; Complete proteome.
SQ SEQUENCE 252 AA; 27130 MW; A401DC1F938C538 CRC64;

Query Match
Best Local Similarity 75.0%; Score 34; DB 17; Length 252;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 EVVPXGMS 9
Db 81 EVIPAGMS 88

RESULT 18
Q96M01 PRELIMINARY; PRT; 290 AA.
ID Q96M01
AC Q96M01;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE CDNA FLJ31891 fis, clone NT2RP7003304, weakly similar to YceA protein
DE homolog ybfQ.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Tashiro H., Yamazaki M., Watanabe K., Kumagai A., Itakura S.,
RA Fukuzumi Y., Fujimori Y., Komiyama M., Sugiyama T., Irie R.,
RA Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,
RA Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K., Wagatsuma M.,
RA Murakawa K., Kanehori K., Takahashi-Fujii A., Oshima A., Sugiyama A.,
RA Kawakami B., Suzuki Y., Sugano S., Nagahari K., Masuho Y., Nagai K.,
RA Isogai T.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AK056453; BAB71188.1; -
DR InterPro: IPR001763; Rhodanese-like.
DR Pfam: PF00581; Rhodanese; 1.
SQ SEQUENCE 290 AA; 32972 MW; 41FB7FB0217C1421 CRC64;

Query Match
Best Local Similarity 65.4%; Score 34; DB 4; Length 290;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 EVVPXGMS 9
Db 35 EEIVPMGIS 43

RESULT 19
Q92MD6 PRELIMINARY; PRT; 387 AA.
ID Q92MD6
AC Q92MD6;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)

```

	01-DEC-2001 (TrEMBLrel. 19, Last sequence update)	
OT	01-MAR-2002 (TrEMBLrel. 20, Last annotation update)	
DE	Putative HIPPURATE hydrolase protein (EC 3.5.1.32).	
EE	HIP01 OR R02690 OR SMC00682.	
NN	Rhizobium meliloti (Sinorhizobium meliloti).	
SN	Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;	
CC	Rhizobiaceae; Sinorhizobium.	
OC	NCBI_TaxID=382;	
RR	[1]	
RN	SEQUENCE FROM N.A.	
RP	STRAIN=1021;	
RC	MEDLINE=21396507; PubMed=11481430;	
RX	Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,	
RA	Bolstad P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,	
RR	Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masuy D.,	
RA	Pohl T., Portetelle D., Puehler A., Fumelle B., Ramsperger U.,	
RR	Renard C., Thebaud P., Vandenberg M., Weidner S., Gallibert F.,	
RA	"Analysis of the chromosome sequence of the legume symbiont	
RT	Sinorhizobium meliloti strain 1021."	
TT	Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).	
RL	EMBL: AL591791; CAC47269.1; -.	
LR	InterPro; IPR002933; Peptidase_M20.	
DR	Fram; PF01546; Peptidase_M20; 1.	
DD	Hydrolase; Complete proteome.	
KW	SEQUENCE 387 AA; 41074 MW; 342763088907A6E3 CRG64;	
QS	Query Match 65.4%; Score 34; DB 16; Length 387;	
	Best Local Similarity 50.0%; Pred. No. 70;	
	Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;	
QY	1 EEVVPXGMYSY 10	
	: : :   :	
DB	367 DEAIPIHGISY 376	
	RESULT 20	
ID	Q98BP5 PRELIMINARY; PRS; 541 AA.	
QD	Q98BP5;	
AC	01-OCT-2001 (TrEMBLrel. 18, Created)	
DT	01-OCT-2001 (TrEMBLrel. 18, Last sequence update)	
DT	01-JUN-2002 (TrEMBLrel. 21, Last annotation update)	
DT	Probable DNA ligase.	
DE	ML5481.	
GN	Rhizobium loti (Mesorhizobium loti).	
OS	Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;	
OC	Phyllobacteriaceae; Mesorhizobium.	
OX	NCBI_TaxID=381;	
OX	[1]	
RN	SEQUENCE FROM N.A.	
RP	STRAIN=MAFF303099;	
RC	MEDLINE=21082930; PubMed=11214968;	
RX	Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,	
RA	Tanabe A., Ishikawa K., Ishikawa A., Kawashima K., Kimura T.,	
RA	Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,	
RA	Mochizuki Y., Nakayama C., Nakazaki N., Shimpo S., Sugimoto M.,	
RA	Takeuchi C., Yamada M., Tabata S.;	
RA	"Complete genome structure of the nitrogen-fixing symbiotic bacterium	
RT	Mesorhizobium loti;"	
TL	DNA Res. 7:331-338(2000).	
RL	EMBL: AP03006; BAB51927.1; -.	
DR	InterPro; IPR000977; DNALigase.	
DR	Fram; PF01068; DNALigase; 1	
DR	PROSITE; PS00697; DNALIGASE_A1; UNKNOWN_1.	
DR	PROSITE; PS50160; DNALIGASE_A3; 1.	
KW	Ligase; Complete proteome.	
QS	SEQUENCE 541 AA; 60645 MW; 2EFEEF705453F28F8 CRC64;	
	Query Match 65.4%; Score 34; DB 16; Length 541;	
	Best Local Similarity 60.0%; Pred. No. 1e+02;	
	Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;	
OY	1 EEVVPXGMYSY 10	

```

QY      3 VVPXGMSYS 11
Db      111 VVPGMAYA 119

RESULT 23
Q8XT05 PRELIMINARY; PRT; 1049 AA.
AC      Q8XT05;
DT      01-MAR-2002 (TREMBlrel. 20, Created)
DT      01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT      01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE      Probable multidrug efflux system transmembrane protein.
GN      MEXD OR RSP0312 OR RS05457.
OS      Ralstonia solanacearum (Pseudomonas solanacearum).
OG      Plasmid megaplasmid.
OC      Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
OC      Ralstonia.
ON      NCBI_TaxID=305;
RN      [1]
RX      STRAIN=GM11000;
RX      MEDLINE=21681879; PubMed=11823852;
RA      Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
RA      Arlat M., Billault A., Brottier P., Camus J.C., Cattolico L.,
RA      Chandler M., Choisme N., Claudel-Renard C., Cunnac S., Demange N.,
RA      Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schiex T.,
RA      Sigulier P., Thebault P., Whalen M., Wincker P., Levy M.,
RA      Weissbach J., Boucher C.A.;
RT      "Genome sequence of the plant pathogen Ralstonia solanacearum.";
RL      Nature 415:497-502(2002).
DR      EMBL; AL646078; CAD17463.1; -
DR      InterPro; IPR001036; Acrflvin_res.
DR      InterPro; IPR000731; HMGR/patch_5TM.
DR      Pfam; PF00873; ACRTRAN; 1.
DR      PRINTS; PR00702; ACRFLVINRP.
DR      TIGRFAMs; TIGR00915; 2A0602; 1.
DR      PROSITE; PS0156; SSD; 1.
KW      Plasmid; Complete proteome.
SQ      SEQUENCE 1049 AA; 111769 MW; CB59674B670089CE CRC64;

Query Match      65.4%; Score 34; DB 16; Length 1049;
Best Local Similarity 75.0%; Pred. No. 2.1e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      4 VVPXGMSYS 11
Db      317 MPAGMSYS 324

RESULT 24
Q96914 PRELIMINARY; PRT; 1499 AA.
AC      Q96914;
DT      01-DEC-2001 (TREMBlrel. 19, Created)
DT      01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT      01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE      Putative aminophospholipid translocase (Aminophospholipid-transporting
DE      ATPase).
GN      ATP10C.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
ON      NCBI_TaxID=9606;
RN      [1]
RX      SEQUENCE FROM N.A.
RX      MEDLINE=2125279; PubMed=11326269;
RA      Meguro M., Kashiwagi A., Mitsuya K., Nakao M., Kondo I., Saitoh S.,
RA      Oshimura M.;
RT      "A novel maternally expressed gene, ATP10C, encodes a putative
RT      aminophospholipid translocase associated with Angelman syndrome.";

```

```

RL      Nat. Genet. 28:19-20(2001).
RN      [2]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=21313119; PubMed=11353404;
RA      Herzog L.B.K., Kim S.J., Cook E.H. Jr., Ledbetter D.H.;
RT      "The human aminophospholipid-transporting ATPase gene ATP10C maps
RT      adjacent to UBE3A and exhibits similar imprinted expression.";
RL      Am. J. Hum. Genet. 68:1501-1505(2001).
DR      EMBL; AB051358; BAB47392.1; -
DR      EMBL; AY029504; AAK33100.1; -
DR      EMBL; AY029487; AAK33100.1; JOINED.
DR      EMBL; AY029488; AAK33100.1; JOINED.
DR      EMBL; AY029489; AAK33100.1; JOINED.
DR      EMBL; AY029490; AAK33100.1; JOINED.
DR      EMBL; AY029491; AAK33100.1; JOINED.
DR      EMBL; AY029492; AAK33100.1; JOINED.
DR      EMBL; AY029493; AAK33100.1; JOINED.
DR      EMBL; AY029494; AAK33100.1; JOINED.
DR      EMBL; AY029495; AAK33100.1; JOINED.
DR      EMBL; AY029496; AAK33100.1; JOINED.
DR      EMBL; AY029497; AAK33100.1; JOINED.
DR      EMBL; AY029498; AAK33100.1; JOINED.
DR      EMBL; AY029499; AAK33100.1; JOINED.
DR      EMBL; AY029500; AAK33100.1; JOINED.
DR      EMBL; AY029501; AAK33100.1; JOINED.
DR      EMBL; AY029502; AAK33100.1; JOINED.
DR      EMBL; AY029503; AAK33100.1; JOINED.
DR      InterPro; IPR001757; ATPase_E1-E2.
DR      InterPro; IPR001064; Crystallin.
DR      Pfam; PF00702; Hydrolase; 1.
DR      PROSITE; PS00154; ATPASE_E1_E2; UNKNOWN.1.
DR      PROSITE; PS00225; CRYSTALLIN_BETAGAMMA; UNKNOWN.1.
SQ      SEQUENCE 1499 AA; 167687 MW; D4996A4D0635A68D CRC64;

Query Match      65.4%; Score 34; DB 4; Length 1499;
Best Local Similarity 72.7%; Pred. No. 3e+02;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 EEVVPXGMSYS 11
Db      469 EEVPRGGSVS 479

RESULT 25
Q8TX62 PRELIMINARY; PRT; 143 AA.
AC      Q8TX62;
DT      01-JUN-2002 (TREMBlrel. 21, Created)
DT      01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT      01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE      Uncharacterized conserved protein.
GN      MK0814.
OS      Methanopyrus kandleri.
OC      Archaea; Euryarchaeota; Methanopyri; Methanopyrales; Methanopyraceae;
OC      Methanopyrus.
ON      NCBI_TaxID=2320;
RN      [1]
RX      SEQUENCE FROM N.A.
RX      STRAIN=AV19 / DSM 6324 / JCM 9639;
RX      MEDLINE=21927647; PubMed=11930014;
RA      Shcherbinina O.V., Shakhova V.V., Makarova K.S., Polushin N.N.,
RA      Natale D.A., Rogozin I.B., Tatusov R.L., Wolf Y.I., Stetter K.O.,
RA      Malykh A.G., Koonin E.V., Kozaykin S.A.;
RT      "The complete genome of hyperthermophile Methanopyrus kandleri AV19
RT      and monophyly of archaeal methanogens.";
RL      Proc. Natl. Acad. Sci. U.S.A. 99:4644-4649(2002).
DR      EMBL; AB010372; AA02027.1; -
KW      Complete proteome.
SQ      SEQUENCE 143 AA; 15734 MW; 4C8B28A1FBEDD0B CRC64;

Query Match      63.5%; Score 33; DB 17; Length 143;

```

Wed Jun 11 15:44:25 2003

Best Local Similarity 60.0%; Pred. No. 40;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEWVPXGMSY 10  
      ||:|||| |  
Db 75 EELVPQGAGY 84

Search completed: June 10, 2003, 13:46:26  
Job time : 25.7857 secs





GenCore version 5.1.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2003, 13:24:45 ; Search time 31.3571 Seconds  
(without alignments)  
46.744 Million cell updates/sec

Title: US-09-909-164-7  
Perfect score: 56  
Sequence: 1 EEVFXGMHYS 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues  
Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 75 summaries

Database :			
A_Geneseq_101002.*			
1:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*		
2:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*		
3:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*		
4:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*		
5:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*		
6:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*		
7:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*		
8:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*		
9:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*		
10:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*		
11:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*		
12:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*		
13:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*		
14:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*		
15:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*		
16:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*		
17:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*		
18:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*		
19:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*		
20:	/SID22/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*		
21:	/SID22/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*		
22:	/SID22/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*		
23:	/SID22/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*		

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	54	96.4	11	23	ABB80523 Hepatitis C virus
2	54	96.4	11	23	ABB80527 Hepatitis C virus
3	54	96.4	11	23	ABB80538 Hepatitis C virus
4	54	96.4	11	23	ABB80560 Hepatitis C virus
5	49	87.5	11	23	ABB80537 Hepatitis C virus
6	49	87.5	11	23	ABB80541 Hepatitis C virus
7	48	85.7	11	23	ABB80546 Hepatitis C virus
8	48	85.7	11	23	ABB80550 Hepatitis C virus
9	48	85.7	11	23	ABB80554 Hepatitis C virus
10	48	85.7	11	23	ABB80555 Hepatitis C virus

11	46	82.1	11	23	ABB80531 Hepatitis C virus
12	46	82.1	11	23	ABB80532 Hepatitis C virus
13	45	80.4	11	23	ABB80521 Hepatitis C virus
14	45	80.4	11	23	ABB80522 Hepatitis C virus
15	45	80.4	11	23	ABB80524 Hepatitis C virus
16	45	80.4	11	23	ABB80525 Hepatitis C virus
17	45	80.4	11	23	ABB80526 Hepatitis C virus
18	45	80.4	11	23	ABB80528 Hepatitis C virus
19	45	80.4	11	23	ABB80529 Hepatitis C virus
20	45	80.4	11	23	ABB80559 Hepatitis C virus
21	45	80.4	11	23	ABB80561 Hepatitis C virus
22	45	80.4	11	23	ABB80562 Hepatitis C virus
23	45	80.4	11	23	ABB80563 Hepatitis C virus
24	45	80.4	11	23	ABB80564 Hepatitis C virus
25	45	80.4	11	23	ABB80565 Hepatitis C virus
26	45	80.4	11	23	ABB80566 Hepatitis C virus
27	45	80.4	11	23	ABB80567 Hepatitis C virus
28	45	80.4	11	23	ABB80568 Hepatitis C virus
29	40	71.4	11	23	ABB80535 Hepatitis C virus
30	40	71.4	11	23	ABB80536 Hepatitis C virus
31	40	71.4	11	23	ABB80538 Hepatitis C virus
32	40	71.4	11	23	ABB80539 Hepatitis C virus
33	40	71.4	11	23	ABB80540 Hepatitis C virus
34	40	71.4	11	23	ABB80542 Hepatitis C virus
35	40	71.4	11	23	ABB80543 Hepatitis C virus
36	39	69.6	11	23	ABB80544 Hepatitis C virus
37	39	69.6	11	23	ABB80545 Hepatitis C virus
38	39	69.6	11	23	ABB80547 Hepatitis C virus
39	39	69.6	11	23	ABB80548 Hepatitis C virus
40	39	69.6	11	23	ABB80549 Hepatitis C virus
41	39	69.6	11	23	ABB80551 Hepatitis C virus
42	39	69.6	11	23	ABB80552 Hepatitis C virus
43	39	69.6	11	23	ABB80553 Hepatitis C virus
44	39	69.6	11	23	ABB80556 Hepatitis C virus
45	39	69.6	11	23	ABB80557 Hepatitis C virus
46	39	69.6	161	21	AA1040335 Human ORF199
47	39	69.6	161	23	ABP07096 Human ORF protein
48	39	69.6	567	22	AAU39521 Propionibacterium
49	39	69.6	1037	23	ABP39949 Staphylococcus epi
50	37	66.1	11	23	ABB80530 Hepatitis C virus
51	37	66.1	11	23	ABB80533 Hepatitis C virus
52	36	64.3	55	23	AAU50964 Hepatitis C virus
53	36	64.3	743	21	AAU99488 Hepatitis C virus
54	35	62.5	20	20	AAU76810 Yeast acyltransfer
55	35	62.5	1022	22	ABG03621 Hepatitis C virus
56	35	62.5	1022	22	ABG05826 Novel human diagno
57	35	62.5	1022	22	ABG08173 Novel human diagno
58	35	62.5	1022	22	ABG08173 Propionibacterium
59	34	60.7	150	22	AAU50013 Drosophila melanog
60	34	60.7	2438	22	ABG59970 Novel human diagno
61	34	60.7	2438	22	ABG21296 Human immune/haema
62	33	58.9	80	22	AAU10294 Human secreted pro
63	33	58.9	92	22	AAU86603 Human albumin fusi
64	33	58.9	200	23	AAU91410 Human secreted pro
65	33	58.9	231	23	ABG65016 Human albumin fusi
66	33	58.9	231	23	AAU91390 Human secreted pro
67	33	58.9	240	23	ABG65018 Human secreted pro
68	33	58.9	240	23	AAU91360 Human polypeptide
69	33	58.9	257	22	AAU10294 Human secreted pro
70	33	58.9	286	23	AAU91409 Murine soluble int
71	33	58.9	379	21	AAU59390 Novel human diagno
72	33	58.9	394	22	ABG07474 Human endocrine po
73	33	58.9	445	22	AAU18340 Novel human diagno
74	33	58.9	449	22	ABG07475 Amino acid sequenc
75	33	58.9	465	20	AAU36985

ALIGNMENTS

RESULT 1  
ABB80523  
ID ABB80523 standard; peptide; 11 AA.



FT FT /note= "Norvalyl carbonyl forming keto-amide linkage with  
FT FT residue 7"  
FT FT Modified-site 8 /note= "Oxymethionine"  
FT FT Modified-site 11 /note= "C-terminal amide"  
FT FT  
XX XX WO200208251-A2.  
XX XX  
XX XX 31-JAN-2002.  
XX XX  
XX XX 19-JUL-2001; 2001WO-US23169.  
XX XX  
XX XX 21-JUL-2000; 2000US-220101P.  
XX XX  
XX XX (CORV-) CORVAS INT INC.  
XX XX  
XX XX Lim-wilby M, Levy OE, Brunck TK;  
XX XX WPI; 2002-361643/39.  
XX XX  
XX XX Novel peptide compound having hepatitis C virus protease inhibitory  
PT activity useful for treating disorders associated with hepatitis C  
PT virus protease  
XX XX  
XX XX Claim 17; Page 65; 69pp; English.  
XX XX  
XX XX The sequence represents a peptide compound of the invention having  
CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
CC invention are alpha-ketoamide peptide analogues. The peptides have  
CC virucide activity, and are useful for treating and in the manufacture of  
CC a medicament to treat disorders associated with HCV protease. A  
CC pharmaceutical composition comprising the peptide as an active ingredient  
CC is useful for treating disorders associated with hepatitis C virus.  
XX XX  
XX XX Sequence 11 AA;  
SQ  
Query Match 96.4%; Score 54; DB 23; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.00045;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 EEVVPXGMHYS 11  
Db 1 EEVVPXGMHYS 11  
RESULT 4  
ABB80560  
ID ABB80560 standard; peptide; 11 AA.  
XX XX  
XX XX ABB80560;  
XX XX  
XX XX 08-OCT-2002 (first entry)  
XX XX  
XX XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #40.  
XX XX  
XX XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
XX virucide.  
XX Synthetic.  
XX XX  
XX XX Key Location/Qualifiers  
FH Key  
FT Modified-site 1 /note= "N-terminal acetyl"  
FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with  
FT residue 7"  
FT Misc-difference 8 /note= "D-form residue"  
FT Modified-site 8 /note= "Oxymethionine"  
FT Misc-difference 9 /note= "D-form residue"  
FT

FT Modified-site 11 /note= "C-terminal amide"  
FT  
XX WO200208251-A2.  
XX  
XX 31-JAN-2002.  
XX  
XX 19-JUL-2001; 2001WO-US23169.  
XX  
XX 21-JUL-2000; 2000US-220101P.  
XX  
XX (CORV-) CORVAS INT INC.  
XX  
XX Lim-wilby M, Levy OE, Brunck TK;  
XX WPI; 2002-361643/39.  
XX  
XX Novel peptide compound having hepatitis C virus protease inhibitory  
PT activity useful for treating disorders associated with hepatitis C  
PT virus protease  
XX  
XX Claim 17; Page 65; 69pp; English.  
XX  
XX The sequence represents a peptide compound of the invention having  
CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
CC invention are alpha-ketoamide peptide analogues. The peptides have  
CC virucide activity, and are useful for treating and in the manufacture of  
CC a medicament to treat disorders associated with HCV protease. A  
CC pharmaceutical composition comprising the peptide as an active ingredient  
CC is useful for treating disorders associated with hepatitis C virus.  
XX  
XX Sequence 11 AA;  
SQ  
Query Match 96.4%; Score 54; DB 23; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.00045;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 EEVVPXGMHYS 11  
Db 1 EEVVPXGMHYS 11  
RESULT 5  
ABB80537  
ID ABB80537 standard; peptide; 11 AA.  
XX XX  
XX XX ABB80537;  
XX XX  
XX XX 08-OCT-2002 (first entry)  
XX XX  
XX XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #17.  
XX XX  
XX XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
XX virucide.  
XX Synthetic.  
XX XX  
XX XX Key Location/Qualifiers  
FH Key  
FT Modified-site 1 /note= "N-terminal acetyl"  
FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with  
FT residue 7"  
FT Misc-difference 9 /note= "D-form residue"  
FT Modified-site 11 /note= "C-terminal amide"  
FT  
XX WO200208251-A2.  
XX  
XX 31-JAN-2002.  
XX  
XX 19-JUL-2001; 2001WO-US23169.

XX PR 21-JUL-2000; 2000US-220101P.  
 XX PA (CORV-) CORVAS INT INC.  
 XX PI Lim-wilby M, Levy OE, Brunck TK;  
 XX DR WPI; 2002-361643/39.  
 XX  
 XX PT Novel peptide compound having hepatitis C virus protease inhibitory  
 PT activity useful for treating disorders associated with hepatitis C  
 PT virus protease -  
 XX  
 XX PS Claim 17; Page 64; 69pp; English.  
 XX  
 CC The sequence represents a peptide compound of the invention having  
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
 CC invention are alpha-ketoamide peptide analogues. The peptides have  
 CC virucide activity, and are useful for treating and in the manufacture of  
 CC a medicament to treat disorders associated with HCV protease. A  
 CC pharmaceutical composition comprising the peptide as an active ingredient  
 CC is useful for treating disorders associated with hepatitis C virus.  
 XX  
 XX SQ Sequence 11 AA;  
 Query Match 87.5%; Score 49; DB 23; Length 11;  
 Best Local Similarity 90.9%; Pred. No. 0.0043;  
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 EEVVPXGMHYS 11  
 ||||| |||  
 DB 1 EEVVPXGMHYS 11  
 ||||| |||

RESULT 6  
 ABB80541  
 ID ABB80541 standard; peptide; 11 AA.  
 XX AC ABB80541;  
 XX  
 XX DT 08-OCT-2002 (first entry)  
 XX  
 DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #21.  
 XX  
 KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
 KW virucide.  
 XX  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Modified-site 1 /note= "N-terminal acetyl"  
 FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with  
 FT residue 7"  
 FT Misc-difference 8 /note= "D-form residue"  
 FT Modified-site 11 /note= "C-terminal amide"  
 FT  
 XX WO200208251-A2.  
 XX  
 XX PN 31-JAN-2002.  
 XX PD  
 XX PF 19-JUL-2001; 2001WO-US23169.  
 XX PR 21-JUL-2000; 2000US-220101P.  
 XX  
 XX PA (CORV-) CORVAS INT INC.  
 XX  
 XX PI Lim-wilby M, Levy OE, Brunck TK;  
 XX DR WPI; 2002-361643/39.  
 XX  
 XX PT Novel peptide compound having hepatitis C virus protease inhibitory  
 PT activity useful for treating disorders associated with hepatitis C  
 PT virus protease -  
 XX  
 XX PS Claim 17; Page 64; 69pp; English.  
 XX  
 CC The sequence represents a peptide compound of the invention having  
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
 CC invention are alpha-ketoamide peptide analogues. The peptides have  
 CC virucide activity, and are useful for treating and in the manufacture of  
 CC a medicament to treat disorders associated with HCV protease. A  
 CC pharmaceutical composition comprising the peptide as an active ingredient  
 CC is useful for treating disorders associated with hepatitis C virus.

XX Novel peptide compound having hepatitis C virus protease inhibitory  
 PT activity useful for treating disorders associated with hepatitis C  
 PT virus protease -  
 XX  
 XX PS Claim 17; Page 65; 69pp; English.  
 XX  
 CC The sequence represents a peptide compound of the invention having  
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
 CC invention are alpha-ketoamide peptide analogues. The peptides have  
 CC virucide activity, and are useful for treating and in the manufacture of  
 CC a medicament to treat disorders associated with HCV protease. A  
 CC pharmaceutical composition comprising the peptide as an active ingredient  
 CC is useful for treating disorders associated with hepatitis C virus.  
 XX  
 XX SQ Sequence 11 AA;  
 Query Match 87.5%; Score 49; DB 23; Length 11;  
 Best Local Similarity 90.9%; Pred. No. 0.0043;  
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 EEVVPXGMHYS 11  
 ||||| |||  
 DB 1 EEVVPXGMHYS 11  
 ||||| |||

RESULT 7  
 ABB80546  
 ID ABB80546 standard; peptide; 11 AA.  
 XX AC ABB80546;  
 XX  
 XX DT 08-OCT-2002 (first entry)  
 XX  
 DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #26.  
 XX  
 KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
 KW virucide.  
 XX  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Modified-site 1 /note= "N-terminal acetyl"  
 FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with  
 FT residue 7"  
 FT Modified-site 11 /note= "C-terminal amide"  
 FT  
 XX WO200208251-A2.  
 XX  
 XX PN 31-JAN-2002.  
 XX PD  
 XX PF 19-JUL-2001; 2001WO-US23169.  
 XX PR 21-JUL-2000; 2000US-220101P.  
 XX  
 XX PA (CORV-) CORVAS INT INC.  
 XX  
 XX PI Lim-wilby M, Levy OE, Brunck TK;  
 XX DR WPI; 2002-361643/39.  
 XX  
 XX PT Novel peptide compound having hepatitis C virus protease inhibitory  
 PT activity useful for treating disorders associated with hepatitis C  
 PT virus protease -  
 XX  
 XX PS Claim 17; Page 65; 69pp; English.  
 XX  
 CC The sequence represents a peptide compound of the invention having  
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
 CC invention are alpha-ketoamide peptide analogues. The peptides have

CC virucide activity, and are useful for treating and in the manufacture of  
 CC a medicament to treat disorders associated with HCV protease. A  
 CC pharmaceutical composition comprising the peptide as an active ingredient  
 CC is useful for treating disorders associated with hepatitis C virus.

XX SQ Sequence 11 AA;

Query Match 85.7%; Score 48; DB 23; Length 11;  
 Best Local Similarity 90.9%; Pred. No. 0.0067;  
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EEVVPXGMHYS 11  
 ||||| |||  
 Db 1 EEVVPXGMHYS 11

RESULT 8  
 ABB80550  
 ID ABB80550 standard; peptide; 11 AA.  
 XX AC ABB80550;  
 XX DT 08-OCT-2002 (first entry)  
 XX DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #30.  
 XX KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
 XX KW virucide.  
 XX OS Synthetic.

XX FH Key Location/Qualifiers  
 XX FT Modified-site 1 /note= "N-terminal acetyl"  
 XX FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with  
 XX FT residue 7"  
 XX FT Misc-difference 9 /note= "D-form residue"  
 XX FT Modified-site 11 /note= "C-terminal amide"  
 XX FT WO200208251-A2.

XX PN 31-JAN-2002.

XX PD 19-JUL-2001; 2001WO-US23169.

XX PF 21-JUL-2000; 2000US-220101P.

XX PR (CORV-) CORVAS INT INC.

XX PA Lim-wilby M, Levy OE, Brunck TK;

XX PI WPI; 2002-361643/39.

XX DR Novel peptide compound having hepatitis C virus protease inhibitory

XX CC activity useful for treating disorders associated with hepatitis C  
 XX PT virus protease  
 XX PS Claim 17; Page 65; 69pp; English.

XX CC The sequence represents a peptide compound of the invention having  
 XX CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
 XX CC invention are alpha-ketoamide peptide analogues. The peptides have  
 XX CC virucide activity, and are useful for treating and in the manufacture of  
 XX CC a medicament to treat disorders associated with HCV protease. A  
 XX CC pharmaceutical composition comprising the peptide as an active ingredient  
 XX CC is useful for treating disorders associated with hepatitis C virus.

XX SQ Sequence 11 AA;

Query Match 85.7%; Score 48; DB 23; Length 11;

Best Local Similarity 90.9%; Pred. No. 0.0067;  
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EEVVPXGMHYS 11  
 ||||| |||  
 Db 1 EEVVPXGMHYS 11

RESULT 9  
 ABB80554  
 ID ABB80554 standard; peptide; 11 AA.

XX AC ABB80554;

XX DT 08-OCT-2002 (first entry)

XX DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #34.

XX KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
 XX KW virucide.

XX OS Synthetic.

XX FH Key Location/Qualifiers

XX FT Modified-site 1 /note= "N-terminal acetyl"

XX FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with  
 XX FT residue 7"

XX FT Misc-difference 8 /note= "D-form residue"

XX FT Modified-site 11 /note= "C-terminal amide"

XX FT WO200208251-A2.

XX PN 31-JAN-2002.

XX PD 19-JUL-2001; 2001WO-US23169.

XX PF 21-JUL-2000; 2000US-220101P.

XX PR (CORV-) CORVAS INT INC.

XX PA Lim-wilby M, Levy OE, Brunck TK;

XX PI WPI; 2002-361643/39.

XX DR Novel peptide compound having hepatitis C virus protease inhibitory

XX CC activity useful for treating disorders associated with hepatitis C  
 XX PT virus protease  
 XX PS Claim 17; Page 65; 69pp; English.

XX CC The sequence represents a peptide compound of the invention having  
 XX CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
 XX CC invention are alpha-ketoamide peptide analogues. The peptides have  
 XX CC virucide activity, and are useful for treating and in the manufacture of  
 XX CC a medicament to treat disorders associated with HCV protease. A  
 XX CC pharmaceutical composition comprising the peptide as an active ingredient  
 XX CC is useful for treating disorders associated with hepatitis C virus.

XX SQ Sequence 11 AA;

Query Match 85.7%; Score 48; DB 23; Length 11;

Best Local Similarity 90.9%; Pred. No. 0.0067;

Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EEVVPXGMHYS 11  
 ||||| |||  
 Db 1 EEVVPXGMHYS 11

## RESULT 10

ABB80555  
ID ABB80555 standard; peptide; 11 AA.

XX AC ABB80555;

XX DT 08-OCT-2002 (first entry)

XX DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #35.

XX DE Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;

XX KW virucide.

XX OS Synthetic.

XX FH Key Location/Qualifiers

XX FT Modified-site 1

XX FT /note= "N-terminal acetyl"

XX FT Modified-site 6

XX FT /note= "Norvalyl carbonyl forming keto-amide linkage with residue 7"

XX FT Misc-difference 8

XX FT /note= "D-form residue"

XX FT Misc-difference 9

XX FT /note= "D-form residue"

XX FT Modified-site 11

XX FT /note= "C-terminal amide"

XX FT WO200208251-A2.

XX PN 31-JAN-2002.

XX PD

XX XX 19-JUL-2001; 2001WO-US23169.

XX XX 21-JUL-2000; 2000US-220101P.

XX XX (CORV-) CORVAS INT INC.

XX XX Lim-wilby M, Levy OE, Brunck TK;

XX XX WPI; 2002-361643/39.

XX Novel peptide compound having hepatitis C virus protease inhibitory activity useful for treating disorders associated with hepatitis C virus protease

XX PS Claim 17; Page 65; 69pp; English.

XX The sequence represents a peptide compound of the invention having hepatitis C virus (HCV) protease inhibitory activity. The peptides of the invention are alpha-ketoamide peptide analogues. The peptides have virucide activity, and are useful for treating and in the manufacture of a medicament to treat disorders associated with HCV protease. A pharmaceutical composition comprising the peptide as an active ingredient is useful for treating disorders associated with hepatitis C virus.

XX SQ Sequence 11 AA;

Query Match 85.7%; Score 48; DB 23; Length 11;

Best Local Similarity 90.9%; Pred. No. 0.0067;

Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXGMHYS 11

|||||

1 EEVVPXGSHYS 11

## RESULT 11

ABB80531

ID ABB80531 standard; peptide; 11 AA.

XX AC ABB80531;

XX XX

DT 08-OCT-2002 (first entry)

XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #11.

XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide; virucide.

XX OS Synthetic.

XX FH Key Location/Qualifiers

XX FT Modified-site 1

XX FT /note= "N-terminal acetyl"

XX FT Modified-site 6

XX FT /note= "Norvalyl carbonyl forming keto-amide linkage with residue 7"

XX FT Modified-site 11

XX FT /note= "C-terminal amide"

XX PN WO200208251-A2.

XX XX 31-JAN-2002.

XX XX 19-JUL-2001; 2001WO-US23169.

XX XX 21-JUL-2000; 2000US-220101P.

XX XX (CORV-) CORVAS INT INC.

XX XX Lim-wilby M, Levy OE, Brunck TK;

XX XX WPI; 2002-361643/39.

XX Novel peptide compound having hepatitis C virus protease inhibitory activity useful for treating disorders associated with hepatitis C virus protease

XX PS Claim 17; Page 64; 69pp; English.

XX The sequence represents a peptide compound of the invention having hepatitis C virus (HCV) protease inhibitory activity. The peptides of the invention are alpha-ketoamide peptide analogues. The peptides have virucide activity, and are useful for treating and in the manufacture of a medicament to treat disorders associated with HCV protease. A pharmaceutical composition comprising the peptide as an active ingredient is useful for treating disorders associated with hepatitis C virus.

XX SQ Sequence 11 AA;

Query Match 82.1%; Score 46; DB 23; Length 11;

Best Local Similarity 90.9%; Pred. No. 0.017;

Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXGMHYS 11

|||||

1 EEVVPXGSHYS 11

## RESULT 12

ABB80532

ID ABB80532 standard; peptide; 11 AA.

XX AC ABB80532;

XX XX

DT 08-OCT-2002 (first entry)

XX DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #12.

XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide; virucide.

XX OS Synthetic.

XX FH Key

Location/Qualifiers



PA (CORV-) CORVAS INT INC.

XX Lim-wilby M, Levy OE, Brunck TK;

XX WPI; 2002-361643/39.

XX Novel peptide compound having hepatitis C virus protease inhibitory  
PT activity useful for treating disorders associated with hepatitis C  
PT virus protease -

XX Claim 17; Page 64; 69pp; English.

XX The sequence represents a peptide compound of the invention having  
CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
CC invention are alpha-ketoamide peptide analogues. The peptides have  
CC virucide activity, and are useful for treating and in the manufacture of  
CC a medicament to treat disorders associated with HCV protease. A  
CC pharmaceutical composition comprising the peptide as an active ingredient  
CC is useful for treating disorders associated with hepatitis C virus.

XX Sequence 11 AA;

Query Match 80.4%; Score 45; DB 23; Length 11;

Best Local Similarity 90.9%; Pred. No. 0.026; Mismatches 0; Gaps 0;

Qy 1 EEVVPXGMHYS 11

Db 1 EEVVPXGMSYS 11

RESULT 15

ABB80524  
ID ABB80524 standard; peptide; 11 AA.

XX ABB80524;

XX 08-OCT-2002 (first entry)

XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #4.

XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
KW virucide.

XX Synthetic.

XX Key Location/Qualifiers

FT Modified-site 1 /note= "N-terminal acetyl"

FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with  
FT residue 7"

FT Misc-difference 9 /note= "D-form residue"

FT Modified-site 11 /note= "C-terminal amide"

XX WO200208251-A2.

XX 31-JAN-2002.

XX 19-JUL-2001; 2001WO-US23169.

XX 21-JUL-2000; 2000US-220101P.

XX (CORV-) CORVAS INT INC.

XX Lim-wilby M, Levy OE, Brunck TK;

XX WPI; 2002-361643/39.

XX Novel peptide compound having hepatitis C virus protease inhibitory  
PT activity useful for treating disorders associated with hepatitis C

PT virus protease -

XX Claim 17; Page 64; 69pp; English.

XX The sequence represents a peptide compound of the invention having  
CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
CC invention are alpha-ketoamide peptide analogues. The peptides have  
CC virucide activity, and are useful for treating and in the manufacture of  
CC a medicament to treat disorders associated with HCV protease. A  
CC pharmaceutical composition comprising the peptide as an active ingredient  
CC is useful for treating disorders associated with hepatitis C virus.

XX Sequence 11 AA;

Query Match 80.4%; Score 45; DB 23; Length 11;

Best Local Similarity 90.9%; Pred. No. 0.026; Mismatches 0; Gaps 0;

Matches 10; Conservative 0; Indels 1; Indels 0;

Qy 1 EEVVPXGMHYS 11

Db 1 EEVVPXGMDYS 11

RESULT 16

ABB80525  
ID ABB80525 standard; peptide; 11 AA.

XX ABB80525;

XX 08-OCT-2002 (first entry)

XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #5.

XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
KW virucide.

XX Synthetic.

XX Key Location/Qualifiers

FT Modified-site 1 /note= "N-terminal acetyl"

FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with  
FT residue 7"

FT Misc-difference 8 /note= "D-form residue"

FT Modified-site 11 /note= "C-terminal amide"

XX WO200208251-A2.

XX 31-JAN-2002.

XX 19-JUL-2001; 2001WO-US23169.

XX 21-JUL-2000; 2000US-220101P.

XX (CORV-) CORVAS INT INC.

XX Lim-wilby M, Levy OE, Brunck TK;

XX WPI; 2002-361643/39.

XX Novel peptide compound having hepatitis C virus protease inhibitory  
PT activity useful for treating disorders associated with hepatitis C  
PT virus protease -

XX Claim 17; Page 64; 69pp; English.

XX The sequence represents a peptide compound of the invention having  
CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
CC invention are alpha-ketoamide peptide analogues. The peptides have  
CC virucide activity, and are useful for treating and in the manufacture of



CC a medicament to treat disorders associated with HCV protease. A  
CC pharmaceutical composition comprising the peptide as an active ingredient  
CC is useful for treating disorders associated with hepatitis C virus.

XX SQ Sequence 11 AA;

Query Match 80.4%; Score 45; DB 23; Length 11;  
Best Local Similarity 90.9%; Pred. No. 0.026; 1; Indels 0; Gaps 0;  
Matches 10; Conservative 0; Mismatches 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXGMHYS 11  
DB 1 EEVVPXGMHYS 11  
|||||

## RESULT 17

ABB80526  
ID ABB80526 standard; peptide; 11 AA.

XX AC ABB80526;  
XX DT 08-OCT-2002 (first entry)  
XX DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #6.  
XX KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
XX KW virucide.

XX OS Synthetic.

XX FH Key Location/Qualifiers

FT Modified-site 1 /note= "N-terminal acetyl"

FT Modified-site 6

FT /note= "Norvalyl carbonyl forming keto-amide linkage with residue 7"

FT Misc-difference 8

FT /note= "D-form residue"

FT Misc-difference 9

FT /note= "D-form residue"

FT Modified-site 11

FT /note= "C-terminal amide"

XX WO200208251-A2.

XX PN 31-JAN-2002.

XX PD 19-JUL-2001; 2001WO-US23169.

XX PF 21-JUL-2000; 2000US-220101P.

XX PR (CORV-) CORVAS INT INC.

XX PA Lim-wilby M, Levy OE, Brunck TK;

XX PI WPI; 2002-361643/39.

XX DR Novel peptide compound having hepatitis C virus protease inhibitory

XX PT activity useful for treating disorders associated with hepatitis C

XX PT virus protease

XX PS Claim 17; Page 64; 69pp; English.

XX The sequence represents a peptide compound of the invention having

XX hepatitis C virus (HCV) protease inhibitory activity. The peptides of the

XX invention are alpha-ketoamide peptide analogues. The peptides have

XX virucide activity, and are useful for treating and in the manufacture of

XX a medicament to treat disorders associated with HCV protease. A

XX pharmaceutical composition comprising the peptide as an active ingredient

XX is useful for treating disorders associated with hepatitis C virus.

XX SQ Sequence 11 AA;

Query Match 80.4%; Score 45; DB 23; Length 11;  
Best Local Similarity 90.9%; Pred. No. 0.026; 1; Indels 0; Gaps 0;  
Matches 10; Conservative 0; Mismatches 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXGMHYS 11  
DB 1 EEVVPXGMHYS 11  
|||||

## RESULT 18.

ABB80528  
ID ABB80528 standard; peptide; 11 AA.

XX AC ABB80528;  
XX DT 08-OCT-2002 (first entry)  
XX DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #8.  
XX KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
XX KW virucide.

XX OS Synthetic.

XX FH Key Location/Qualifiers

FT Modified-site 1 /note= "N-terminal acetyl"

FT Modified-site 6

FT /note= "Norvalyl carbonyl forming keto-amide linkage with residue 7"

FT Misc-difference 8

FT /note= "D-form residue"

FT Modified-site 11

FT /note= "C-terminal amide"

XX WO200208251-A2.

XX PN 31-JAN-2002.

XX PD 19-JUL-2001; 2001WO-US23169.

XX PF 21-JUL-2000; 2000US-220101P.

XX PR (CORV-) CORVAS INT INC.

XX PI Lim-wilby M, Levy OE, Brunck TK;

XX WPI; 2002-361643/39.

XX Novel peptide compound having hepatitis C virus protease inhibitory

XX activity useful for treating disorders associated with hepatitis C

XX virus protease

XX PS Claim 17; Page 64; 69pp; English.

XX The sequence represents a peptide compound of the invention having

XX hepatitis C virus (HCV) protease inhibitory activity. The peptides of the

XX invention are alpha-ketoamide peptide analogues. The peptides have

XX virucide activity, and are useful for treating and in the manufacture of

XX a medicament to treat disorders associated with HCV protease. A

XX pharmaceutical composition comprising the peptide as an active ingredient

XX is useful for treating disorders associated with hepatitis C virus.

XX SQ Sequence 11 AA;

Query Match 80.4%; Score 45; DB 23; Length 11;  
Best Local Similarity 90.9%; Pred. No. 0.026;  
Matches 10; Conservative 0; Mismatches 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXGMHYS 11  
DB 1 EEVVPXGMHYS 11  
|||||

## RESULT 19

ABB80529  
ID ABB80529 standard; peptide; 11 AA.

AC ABB80529;

DT 08-OCT-2002 (first entry)

DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #9.

KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;

XX virucide.

OS Synthetic.

Key Location/Qualifiers

Modified-site 1 /note= "N-terminal acetyl"

Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with

residue 7"

Misc-difference 8 /note= "D-form residue"

Misc-difference 9 /note= "D-form residue"

Modified-site 11 /note= "C-terminal amide"

WO200208251-A2.

31-JAN-2002.

19-JUL-2001; 2001WO-US23169.

21-JUL-2000; 2000US-220101P.

(CORV-) CORVAS INT INC.

Lim-wilby M, Levy OE, Brunck TK;

WPI; 2002-361643/39.

Novel peptide compound having hepatitis C virus protease inhibitory

activity useful for treating disorders associated with hepatitis C

virus protease

Claim 17; Page 64; 69pp; English.

The sequence represents a peptide compound of the invention having

hepatitis C virus (HCV) protease inhibitory activity. The peptides of the

invention are alpha-ketoamide peptide analogues. The peptides have

virucide activity, and are useful for treating and in the manufacture of

a medicament to treat disorders associated with HCV protease. A

pharmaceutical composition comprising the peptide as an active ingredient

is useful for treating disorders associated with hepatitis C virus.

Query Match 80.4%; Score 45; DB 23; Length 11;

Best Local Similarity 90.9%; Pred. No. 0.026;

Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXGMHYS 11

Db 1 EEVVPXGMHYS 11

RESULT 20

ABB80559

ID ABB80559 standard; peptide; 11 AA.

XX ABB80559;

AC ABB80559;

XX

DT

08-OCT-2002 (first entry)

XX

DE

Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #39.

XX

KW

Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;

XX

KW

virucide.

XX

OS

Synthetic.

XX

Key

Location/Qualifiers

Modified-site 1 /note= "N-terminal acetyl"

Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with

residue 7"

Misc-difference 8 /note= "D-form residue"

Modified-site 8 /note= "D-form residue"

Modified-site 11 /note= "Oxymethionine"

Modified-site 11 /note= "C-terminal amide"

WO200208251-A2.

31-JAN-2002.

19-JUL-2001; 2001WO-US23169.

21-JUL-2000; 2000US-220101P.

(CORV-) CORVAS INT INC.

Lim-wilby M, Levy OE, Brunck TK;

WPI; 2002-361643/39.

Novel peptide compound having hepatitis C virus protease inhibitory

activity useful for treating disorders associated with hepatitis C

virus protease

Claim 17; Page 65; 69pp; English.

The sequence represents a peptide compound of the invention having

hepatitis C virus (HCV) protease inhibitory activity. The peptides of the

invention are alpha-ketoamide peptide analogues. The peptides have

virucide activity, and are useful for treating and in the manufacture of

a medicament to treat disorders associated with HCV protease. A

pharmaceutical composition comprising the peptide as an active ingredient

is useful for treating disorders associated with hepatitis C virus.

Query Match 80.4%; Score 45; DB 23; Length 11;

Best Local Similarity 90.9%; Pred. No. 0.026;

Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXGMHYS 11

Db 1 EEVVPXGMHYS 11

RESULT 21

ABB80561

ID ABB80561 standard; peptide; 11 AA.

XX

AC

ABB80561;

XX

AC

ABB80561;

XX

DT

08-OCT-2002 (first entry)

XX

DE

Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #41.

XX

KW

Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;

KW virucide.  
XX  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT Modified-site 1  
FT Modified-site 6 /note= "N-terminal acetyl"  
FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with residue 7"  
FT  
FT Misc-difference 8  
FT Modified-site 8 /note= "D-form residue"  
FT Modified-site 8 /note= "Oxymethionine"  
FT Modified-site 11  
FT Modified-site 11 /note= "C-terminal amide"  
XX WO200208251-A2.  
XX  
XX 31-JAN-2002.  
XX  
XX 19-JUL-2001; 2001WO-US23169.  
XX  
XX 21-JUL-2000; 2000US-220101P.  
XX (CORV-) CORVAS INT INC.  
XX Lim-wilby M, Levy OE, Brunck TK;  
XX WPI; 2002-361643/39.  
XX Novel peptide compound having hepatitis C virus protease inhibitory activity useful for treating disorders associated with hepatitis C virus protease  
XX  
XX Claim 17; Page 65; 69pp; English.  
XX The sequence represents a peptide compound of the invention having hepatitis C virus (HCV) protease inhibitory activity. The peptides of the invention are alpha-ketoamide peptide analogues. The peptides have virucide activity, and are useful for treating and in the manufacture of a medicament to treat disorders associated with HCV protease. A pharmaceutical composition comprising the peptide as an active ingredient is useful for treating disorders associated with hepatitis C virus.  
XX  
XX Sequence 11 AA;  
SQ  
Query Match 80.4%; Score 45; DB 23; Length 11;  
Best Local Similarity 90.9%; Pred. No. 0.026;  
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 EEVVPXGMHYS 11  
DB 1 EEVVPXGMHYS 11  
RESULT 22  
ABB80562  
ID ABB80562 standard; peptide; 11 AA.  
XX  
XX ABB80562;  
AC  
DT 08-OCT-2002 (first entry)  
DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #42.  
XX  
XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide; virucide.  
XX Synthetic.  
XX Key Location/Qualifiers  
FH Modified-site 1  
FT Modified-site 6 /note= "N-terminal acetyl"  
FT Modified-site 6 /note= "Valyl carbonyl forming keto-amide linkage with residue 7"  
FT

FT Modified-site 6 /note= "N-terminal acetyl"  
FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with residue 7"  
FT  
FT Misc-difference 8  
FT Modified-site 8 /note= "D-form residue"  
FT Modified-site 8 /note= "Oxymethionine"  
FT Misc-difference 9  
FT Modified-site 11  
FT Modified-site 11 /note= "C-terminal amide"  
XX WO200208251-A2.  
XX 31-JAN-2002.  
XX 19-JUL-2001; 2001WO-US23169.  
XX 21-JUL-2000; 2000US-220101P.  
XX (CORV-) CORVAS INT INC.  
XX Lim-wilby M, Levy OE, Brunck TK;  
XX WPI; 2002-361643/39.  
XX Novel peptide compound having hepatitis C virus protease inhibitory activity useful for treating disorders associated with hepatitis C virus protease  
XX  
XX Claim 17; Page 65; 69pp; English.  
XX The sequence represents a peptide compound of the invention having hepatitis C virus (HCV) protease inhibitory activity. The peptides of the invention are alpha-ketoamide peptide analogues. The peptides have virucide activity, and are useful for treating and in the manufacture of a medicament to treat disorders associated with HCV protease. A pharmaceutical composition comprising the peptide as an active ingredient is useful for treating disorders associated with hepatitis C virus.  
XX  
XX Sequence 11 AA;  
SQ  
Query Match 80.4%; Score 45; DB 23; Length 11;  
Best Local Similarity 90.9%; Pred. No. 0.026;  
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 EEVVPXGMHYS 11  
DB 1 EEVVPXGMHYS 11  
RESULT 23  
ABB80563  
ID ABB80563 standard; peptide; 11 AA.  
XX  
XX ABB80563;  
AC  
DT 08-OCT-2002 (first entry)  
DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #43.  
XX  
XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide; virucide.  
XX Synthetic.  
XX Key Location/Qualifiers  
FH Modified-site 1  
FT Modified-site 6 /note= "N-terminal acetyl"  
FT Modified-site 6 /note= "Valyl carbonyl forming keto-amide linkage with residue 7"  
FT

FT Modified-site 11 /note= "C-terminal amide"  
XX WO200208251-A2.  
PN 31-JAN-2002.  
XX  
XX  
XX  
XX  
XX 19-JUL-2001; 2001WO-US23169.  
XX  
XX 21-JUL-2000; 2000US-220101P.  
XX  
XX (CORV-) CORVAS INT INC.  
XX  
XX Lim-wilby M, Levy OE, Brunck TK;  
XX WPI; 2002-361643/39.  
XX  
XX Novel peptide compound having hepatitis C virus protease inhibitory  
PT activity useful for treating disorders associated with hepatitis C  
PT virus protease  
XX  
XX Claim 17; Page 65; 69pp; English.  
XX  
XX The sequence represents a peptide compound of the invention having  
CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
CC invention are alpha-ketoamide peptide analogues. The peptides have  
CC virucide activity, and are useful for treating and in the manufacture of  
CC a medicament to treat disorders associated with HCV protease. A  
CC pharmaceutical composition comprising the peptide as an active ingredient  
CC is useful for treating disorders associated with hepatitis C virus.  
XX  
XX Sequence 11 AA;  
XX

Query Match 80.4%; Score 45; DB 23; Length 11;  
Best Local Similarity 90.9%; Pred. No. 0.026;  
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 EEVVPXGMHYS 11  
DB 1 EEVVPXGMHYS 11  
1111111111

RESULT 24  
ABB80564  
ID ABB80564 standard; peptide; 11 AA.  
XX  
XX ABB80564;  
XX  
XX 08-OCT-2002 (first entry)  
XX  
XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #44.  
XX  
XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
XX virucide.  
XX Synthetic.  
XX  
XX Key Location/Qualifiers  
FT Modified-site 1 /note= "N-terminal acetyl"  
FT Modified-site 6  
FT Modified-site /note= "Leucyl carbonyl forming keto-amide linkage with  
FT residue 7"  
FT Modified-site 11 /note= "C-terminal amide"  
FT  
XX WO200208251-A2.  
XX  
XX 31-JAN-2002.  
XX  
XX 19-JUL-2001; 2001WO-US23169.  
XX  
XX 21-JUL-2000; 2000US-220101P.  
XX  
XX (CORV-) CORVAS INT INC.  
XX  
XX Lim-wilby M, Levy OE, Brunck TK;  
XX WPI; 2002-361643/39.  
XX  
XX Novel peptide compound having hepatitis C virus protease inhibitory  
PT activity useful for treating disorders associated with hepatitis C  
PT virus protease

XX (CORV-) CORVAS INT INC.  
XX Lim-wilby M, Levy OE, Brunck TK;  
XX WPI; 2002-361643/39.  
XX  
XX Novel peptide compound having hepatitis C virus protease inhibitory  
PT activity useful for treating disorders associated with hepatitis C  
PT virus protease  
XX  
XX Claim 17; Page 65; 69pp; English.  
XX  
XX The sequence represents a peptide compound of the invention having  
CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
CC invention are alpha-ketoamide peptide analogues. The peptides have  
CC virucide activity, and are useful for treating and in the manufacture of  
CC a medicament to treat disorders associated with HCV protease. A  
CC pharmaceutical composition comprising the peptide as an active ingredient  
CC is useful for treating disorders associated with hepatitis C virus.  
XX  
XX Sequence 11 AA;  
XX

Query Match 80.4%; Score 45; DB 23; Length 11;  
Best Local Similarity 90.9%; Pred. No. 0.026;  
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 EEVVPXGMHYS 11  
DB 1 EEVVPXGMHYS 11  
1111111111

RESULT 25  
ABB80565  
ID ABB80565 standard; peptide; 11 AA.  
XX  
XX ABB80565;  
XX  
XX 08-OCT-2002 (first entry)  
XX  
XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #45.  
XX  
XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
XX virucide.  
XX Synthetic.  
XX  
XX Key Location/Qualifiers  
FT Modified-site 1 /note= "N-terminal acetyl"  
FT Modified-site 6  
FT Modified-site /note= "Norleucyl carbonyl forming keto-amide linkage  
FT with residue 7"  
FT Modified-site 11 /note= "C-terminal amide"  
FT  
XX WO200208251-A2.  
XX  
XX 31-JAN-2002.  
XX  
XX 19-JUL-2001; 2001WO-US23169.  
XX  
XX 21-JUL-2000; 2000US-220101P.  
XX  
XX (CORV-) CORVAS INT INC.  
XX  
XX Lim-wilby M, Levy OE, Brunck TK;  
XX WPI; 2002-361643/39.  
XX  
XX Novel peptide compound having hepatitis C virus protease inhibitory  
PT activity useful for treating disorders associated with hepatitis C  
PT virus protease

XX Claim 17; Page 65; 69pp; English.  
 PS  
 XX  
 CC The sequence represents a peptide compound of the invention having  
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
 CC invention are alpha-ketoamide peptide analogues. The peptides have  
 CC virucide activity, and are useful for treating and in the manufacture of  
 CC a medicament to treat disorders associated with HCV protease. A  
 CC pharmaceutical composition comprising the peptide as an active ingredient  
 CC is useful for treating disorders associated with hepatitis C virus.

XX Sequence 11 AA;  
 SQ  
 Query Match 80.4%; Score 45; DB 23; Length 11;  
 Best Local Similarity 90.9%; Pred. No. 0.026;  
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXGMHYS 11  
 |||||  
 Db 1 EEVVPXGMSYS 11

Search completed: June 10, 2003, 13:39:06  
 Job time : 31.3571 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2003, 13:31:45 ; Search time 9.64286 Seconds  
(without alignments)  
33.564 Million cell updates/sec

Title: us-09-909-164-7  
Perfect score: 56  
Sequence: 1 EEVVPXGMHYS 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues  
Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 75 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/ptodata/1/iaa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/1/iaa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/1/iaa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/1/iaa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS.COMB.pep:\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	39	69.6	1037	4	US-09-134-001C-4794
2	34	60.7	600	2	US-08-821-119-19
3	34	60.7	600	2	US-08-821-118-2
4	33	58.9	747	4	US-09-724-864-36
5	32	57.1	70	4	US-09-134-001C-3950
6	32	57.1	102	2	US-08-580-988A-23
7	32	57.1	126	2	US-08-879-995A-3
8	32	57.1	126	3	US-09-215-096-3
9	32	57.1	152	2	US-08-460-694-4
10	32	57.1	152	3	US-08-460-744-4
11	32	57.1	152	3	US-07-667-711B-4
12	32	57.1	173	1	US-08-193-977-7
13	32	57.1	189	2	US-08-464-517-21
14	32	57.1	189	2	US-08-246-361A-21
15	32	57.1	189	3	US-08-463-772-21
16	32	57.1	189	5	PCT-US93-05000-21
17	32	57.1	236	2	US-08-464-517-22
18	32	57.1	236	2	US-08-246-361A-22
19	32	57.1	236	2	US-08-463-772-22
20	32	57.1	236	5	PCT-US93-05000-22
21	32	57.1	280	3	US-08-464-517-6
22	32	57.1	280	3	US-08-463-772-6
23	32	57.1	289	2	US-08-246-361A-4
24	32	57.1	289	5	PCT-US93-05000-4
25	32	57.1	291	5	PCT-US93-05000-6
26	32	57.1	292	2	US-08-464-517-23
27	32	57.1	292	2	US-08-246-361A-6

28	57.1	292	2	US-08-246-361A-23	Sequence 23, Appl
29	57.1	292	3	US-08-463-772-23	Sequence 23, Appl
30	57.1	292	5	PCT-US93-05000-23	Sequence 23, Appl
31	57.1	295	1	US-07-947-120-8	Sequence 8, Appl
32	57.1	295	1	US-08-472-893A-8	Sequence 8, Appl
33	57.1	295	2	US-08-460-694-2	Sequence 2, Appl
34	57.1	295	2	US-08-464-517-19	Sequence 19, Appl
35	57.1	295	2	US-08-464-517-20	Sequence 20, Appl
36	57.1	295	2	US-08-246-361A-19	Sequence 19, Appl
37	57.1	295	2	US-08-246-361A-20	Sequence 20, Appl
38	57.1	295	3	US-08-463-772-19	Sequence 19, Appl
39	57.1	295	3	US-08-463-772-20	Sequence 20, Appl
40	57.1	295	3	US-08-460-744-2	Sequence 2, Appl
41	57.1	295	3	US-07-667-711B-2	Sequence 2, Appl
42	57.1	295	3	US-08-947-492-8	Sequence 8, Appl
43	57.1	295	5	PCT-US93-05000-2	Sequence 2, Appl
44	57.1	295	5	PCT-US93-05000-19	Sequence 19, Appl
45	57.1	295	5	PCT-US93-05000-20	Sequence 20, Appl
46	57.1	309	3	US-08-464-517-4	Sequence 4, Appl
47	57.1	309	3	US-08-463-772-4	Sequence 4, Appl
48	57.1	615	2	US-08-663-566A-17	Sequence 17, Appl
49	57.1	615	2	US-08-023-610-17	Sequence 17, Appl
50	57.1	615	2	US-08-288-065A-17	Sequence 17, Appl
51	57.1	615	2	US-08-362-240A-17	Sequence 17, Appl
52	57.1	615	5	PCT-US95-10245-17	Sequence 17, Appl
53	57.1	618	2	US-08-770-761A-3	Sequence 3, Appl
54	57.1	647	2	US-08-770-761A-8	Sequence 8, Appl
55	57.1	660	2	US-08-770-761A-2	Sequence 2, Appl
56	57.1	662	2	US-08-770-761A-5	Sequence 5, Appl
57	57.1	705	2	US-08-770-761A-7	Sequence 7, Appl
58	57.1	819	2	US-08-464-517-7	Sequence 7, Appl
59	57.1	819	2	US-08-246-361A-7	Sequence 7, Appl
60	57.1	819	3	US-08-463-772-7	Sequence 7, Appl
61	57.1	819	5	PCT-US93-05000-7	Sequence 7, Appl
62	56.2	501	2	US-08-480-095-31	Sequence 31, Appl
63	55.4	59	4	US-08-963-851-14	Sequence 14, Appl
64	55.4	65	6	517197-51	Patent No. 517197
65	55.4	410	6	517197-1	Patent No. 517197
66	55.4	447	4	US-08-961-083-182	Sequence 182, Appl
67	55.4	502	4	US-03-342-647-4	Sequence 4, Appl
68	55.4	529	4	US-09-240-639-4	Sequence 4, Appl
69	55.4	622	2	US-08-459-146-2	Sequence 2, Appl
70	55.4	622	2	US-08-459-065-2	Sequence 2, Appl
71	55.4	630	4	US-09-342-647-2	Sequence 2, Appl
72	55.4	667	4	US-09-342-647-28	Sequence 28, Appl
73	55.4	738	1	US-08-530-010-3	Sequence 3, Appl
74	55.4	738	1	US-08-530-010-5	Sequence 5, Appl
75	55.4	738	1	US-08-530-010-7	Sequence 7, Appl

## ALIGNMENTS

RESULT 1  
US-09-134-001C-4794  
; Sequence 4794, Application US/09134001C  
; Patent No. 6380370  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: GTC-007  
; CURRENT APPLICATION NUMBER: US/09/134,001C  
; CURRENT FILING DATE: 1998-08-13  
; PRIOR APPLICATION NUMBER: US 60/064,964  
; PRIOR FILING DATE: 1997-11-08  
; PRIOR APPLICATION NUMBER: US 60/055,779  
; PRIOR FILING DATE: 1997-08-14  
; NUMBER OF SEQ ID NOS: 5674  
; SEQ ID NO 4794  
; LENGTH: 1037  
; TYPE: PRT  
; ORGANISM: Staphylococcus epidermidis

US-09-134-001C-4794

Query Match 59.6%; Score 39; DB 4; Length 1037;  
Best Local Similarity 63.6%; Pred. No. 16;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEPVPGMHYS 11  
:|||||:|||||  
Db 199 KEVVSNGHYS 209

## RESULT 2

US-08-821-119-19  
; Sequence 19, Application US/08821119  
; Patent No. 5821104  
; GENERAL INFORMATION:  
; APPLICANT: Holm, Kaj Andre  
; APPLICANT: Rasmussen, Grethe  
; APPLICANT: Halkier, Torben  
; APPLICANT: Lehmebeck, Jan  
; TITLE OF INVENTION: Tripeptidyl Aminopeptidase  
; NUMBER OF SEQUENCES: 23  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: No. 5821104o No. 5821104disk of No. 5821104th America, Inc.  
; STREET: 405 Lexington Avenue  
; CITY: New York  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10174

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/821.119  
; FILING DATE: 19-MAR-1997  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lambiris, Elias J  
; REGISTRATION NUMBER: 33,728  
; REFERENCE/DOCKET NUMBER: 4107.204-US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-867-0123  
; TELEFAX: 212-878-9655  
; TELEX:

INFORMATION FOR SEQ ID NO: 19:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 600 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; FRAGMENT TYPE: internal  
US-08-821-119-19

Query Match 60.7%; Score 34; DB 2; Length 600;  
Best Local Similarity 75.0%; Pred. No. 82;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 VPXGMHYS 11  
:|||||  
Db 31 VPRGWHYS 38

## RESULT 3

US-08-821-118-2  
; Sequence 2, Application US/08821118  
; Patent No. 5989889  
; GENERAL INFORMATION:  
; APPLICANT: Rev. Michael  
; APPLICANT: Golightly, Elizabeth  
; TITLE OF INVENTION: NUCLEIC ACIDS ENCODING POLYPEPTIDES

TITLE OF INVENTION: HAVING TRIPEPTIDE AMINOPEPTIDASE  
; TITLE OF INVENTION: ACTIVITY  
; NUMBER OF SEQUENCES: 7  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: No. 5989889o No. 5989889disk of No. 5989889th America, Inc.  
; STREET: 405 Lexington Avenue  
; CITY: New York  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10174

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/821.118  
; FILING DATE: 19-MAR-1997  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lambiris, Elias J  
; REGISTRATION NUMBER: 33,728  
; REFERENCE/DOCKET NUMBER: 4107.400-US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-867-0123  
; TELEFAX: 212-878-9655  
; TELEX:

INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 600 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; FRAGMENT TYPE: internal  
US-08-821-118-2

Query Match 60.7%; Score 34; DB 2; Length 600;  
Best Local Similarity 75.0%; Pred. No. 82;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 VPXGMHYS 11  
:|||||  
Db 31 VPRGWHYS 38

## RESULT 4

US-09-724-864-36  
; Sequence 36, Application US/09724864  
; Patent No. 6380362  
; GENERAL INFORMATION:  
; APPLICANT: Watson, James D  
; APPLICANT: Murison, James G.  
; TITLE OF INVENTION: Polynucleotides, polypeptides expressed  
; TITLE OF INVENTION: by the polynucleotides and methods for their use.  
; FILE REFERENCE: 11000.105001  
; CURRENT APPLICATION NUMBER: US/09/724.864  
; CURRENT FILING DATE: 2000-11-28  
; PRIOR APPLICATION NUMBER: U.S. No. 6380362 60/171,678  
; PRIOR FILING DATE: 1999-12-23  
; NUMBER OF SEQ ID NOS: 72  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 36  
; LENGTH: 747  
; TYPE: PRT  
; ORGANISM: Rat  
US-09-724-864-36

Query Match 58.9%; Score 33; DB 4; Length 747;  
Best Local Similarity 71.4%; Pred. No. 1.6e+02;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 PXGMHYS 11



us-09-909-164-7-rai

Wed Jun 11 15:44:28 2003

LENGTH: 102 amino acids  
 TYPE: amino acid  
 STRANDEDNESS:  
 TOPOLOGY: linear  
 MOLECULE TYPE:  
 DESCRIPTION: protein  
 HYPOTHETICAL: no  
 ANTI-SENSE: no  
 FRAGMENT TYPE: Internal  
 ORIGINAL SOURCE:  
 US-08-580-988A-23

Query Match 57.1%; Score 32; DB 2; Length 102;  
 Best Local Similarity 60.0%; Pred. No. 29;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMHY 10  
 ||| | | |  
 Db 24 EEVFPPLAMNY 33

RESULT 7  
 US-08-879-995A-3  
 ; Sequence 3, Application US/08879995A  
 ; Patent No. 5985606  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Hillman, Jennifer L.  
 ; APPLICANT: Lal, Preeti  
 ; APPLICANT: Kaser, Matthew R.  
 ; TITLE OF INVENTION: HUMAN PREPROTACHYKININ B  
 ; NUMBER OF SEQUENCES: 4  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
 ; STREET: 3174 Porter Drive  
 ; CITY: Palo Alto  
 ; STATE: CA  
 ; COUNTRY: USA  
 ; ZIP: 94304  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette  
 ; COMPUTER: IBM Compatible  
 ; OPERATING SYSTEM: DOS  
 ; SOFTWARE: FASTSEQ for Windows Version 2.0  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/879,995A  
 ; FILING DATE: Herewith  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER:  
 ; FILING DATE:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Billings, Lucy J.  
 ; REGISTRATION NUMBER: 36,749  
 ; REFERENCE/DOCKET NUMBER: PF-0326 US  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 415-855-0555  
 ; TELEFAX: 415-845-4166  
 ; TELEX:  
 ; INFORMATION FOR SEQ ID NO: 3:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 126 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; IMMEDIATE SOURCE:  
 ; LIBRARY: GenBank  
 ; CLONE: 163590  
 ; US-08-879-995A-3

Query Match 57.1%; Score 32; DB 2; Length 126;  
 Best Local Similarity 66.7%; Pred. No. 36;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Db 627 PGLHYS 633  
 | | | | |

RESULT 5  
 US-09-134-001C-3950  
 ; Sequence 3950, Application US/09134001C  
 ; Patent No. 6380370  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lynn Doucette-Stamm et al  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
 ; FILE REFERENCE: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS  
 ; CURRENT APPLICATION NUMBER: US/09/134,001C  
 ; CURRENT FILING DATE: 1998-08-13  
 ; PRIOR APPLICATION NUMBER: US 60/064,964  
 ; PRIOR FILING DATE: 1997-11-08  
 ; PRIOR APPLICATION NUMBER: US 60/055,779  
 ; PRIOR FILING DATE: 1997-08-14  
 ; NUMBER OF SEQ ID NOS: 5674  
 ; SEQ ID NO 3950  
 ; LENGTH: 70  
 ; TYPE: PRT  
 ; ORGANISM: Staphylococcus epidermidis  
 ; US-09-134-001C-3950

Query Match 57.1%; Score 32; DB 4; Length 70;  
 Best Local Similarity 62.5%; Pred. No. 19;  
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 VPXGMHYS 11  
 : | | | | |  
 Db 36 MPKGFHYS 43

RESULT 6  
 US-08-580-988A-23  
 ; Sequence 23, Application US/08580988A  
 ; Patent No. 5856161  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Aggarwal et al.  
 ; TITLE OF INVENTION: Tumor Necrosis Factor  
 ; TITLE OF INVENTION: Receptor-1-Associated Protein Kinase And Methods  
 ; TITLE OF INVENTION: For Its Use  
 ; NUMBER OF SEQUENCES: 27  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Dr. Benjamin A. Adler  
 ; STREET: 8011 Candle Lane  
 ; CITY: Houston  
 ; STATE: Texas  
 ; COUNTRY: USA  
 ; ZIP: 77071  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: 1.44 Mb floppy disk  
 ; COMPUTER: Apple Macintosh  
 ; OPERATING SYSTEM: Macintosh  
 ; SOFTWARE: Microsoft Word for Macintosh  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/580,988A  
 ; FILING DATE: January 3, 1996  
 ; CLASSIFICATION: 424  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER:  
 ; FILING DATE:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Benjamin Aaron Adler, Ph.D., J.D.  
 ; REGISTRATION NUMBER: 35,423  
 ; REFERENCE/DOCKET NUMBER: D5721CIP2  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 713-777-2321  
 ; TELEFAX: 713-777-6908  
 ; INFORMATION FOR SEQ ID NO: 23:  
 ; SEQUENCE CHARACTERISTICS:

QY 1 EEVVPXGMH 9  
 I:| | | | |  
 Db 28 EQVVPGGH 36

RESULT 8  
 US-09-215-096-3  
 ; Sequence 3, Application US/09215096  
 ; Patent No. 6008194  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Hillman, Jennifer L.  
 ; APPLICANT: Lal, Preeti  
 ; APPLICANT: Kaser, Matthew R.  
 ; TITLE OF INVENTION: HUMAN PREPROTACHYKININ B  
 ; NUMBER OF SEQUENCES: 4  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
 ; STREET: 3174 Porter Drive  
 ; CITY: Palo Alto  
 ; STATE: CA  
 ; COUNTRY: USA  
 ; ZIP: 94304  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette  
 ; COMPUTER: IBM Compatible  
 ; OPERATING SYSTEM: DOS  
 ; SOFTWARE: FastSeq for Windows Version 2.0  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/215,096  
 ; FILING DATE:  
 ; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/879,995  
 ; FILING DATE:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Billings, Lucy J.  
 ; REGISTRATION NUMBER: 36,749  
 ; REFERENCE/DOCKET NUMBER: PF-0326 US  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 415-855-0555  
 ; TELEFAX: 415-845-4166  
 ; TELEX:  
 ; INFORMATION FOR SEQ ID NO: 3:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 126 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; IMMEDIATE SOURCE:  
 ; LIBRARY: GenBank  
 ; CLONE: 163590  
 ; US-09-215-096-3

Query Match 57.1%; Score 32; DB 3; Length 126;  
 Best Local Similarity 66.7%; Pred. No. 36;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMH 9  
 I:| | | | |  
 Db 28 EQVVPGGH 36

RESULT 9  
 US-08-460-694-4  
 ; Sequence 4, Application US/08460694  
 ; Patent No. 5858655  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Arnold, Andrew  
 ; TITLE OF INVENTION: PRADI Cyclin and its cdNA  
 ; NUMBER OF SEQUENCES: 8  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
 ; STREET: 1100 New York Avenue, N.W., Suite 600

; CITY: Washington  
 ; STATE: DC  
 ; COUNTRY: USA  
 ; ZIP: 20005  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/460,694  
 ; FILING DATE: 02-JUN-1995  
 ; CLASSIFICATION: 435  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: McConathy, Evelyn H.  
 ; REGISTRATION NUMBER: 35,279  
 ; REFERENCE/DOCKET NUMBER: 0609.4070002  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 202-371-2600  
 ; TELEFAX: 202-371-2540  
 ; INFORMATION FOR SEQ ID NO: 4:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 152 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: not relevant  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: peptide  
 ; US-08-460-694-4

Query Match 57.1%; Score 32; DB 2; Length 152;  
 Best Local Similarity 60.0%; Pred. No. 44;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMH 10  
 I:| | | | |  
 Db 20 EEVPLAMNY 29

RESULT 10  
 US-08-460-744-4  
 ; Sequence 4, Application US/08460744  
 ; Patent No. 6107541  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Arnold, Andrew  
 ; TITLE OF INVENTION: PRADI Cyclin and its cdNA  
 ; NUMBER OF SEQUENCES: 8  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
 ; STREET: 1100 New York Avenue, N.W., Suite 600  
 ; CITY: Washington  
 ; STATE: DC  
 ; COUNTRY: USA  
 ; ZIP: 20005  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/460,744  
 ; FILING DATE: 02-JUN-1995  
 ; CLASSIFICATION: 435  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: McConathy, Evelyn H.  
 ; REGISTRATION NUMBER: 35,279  
 ; REFERENCE/DOCKET NUMBER: 0609.4070005  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 202-371-2600  
 ; TELEFAX: 202-371-2540  
 ; INFORMATION FOR SEQ ID NO: 4:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 152 amino acids  
 ; TYPE: amino acid

STRANDEDNESS: not relevant  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-460-744-4

Query Match 57.1%; Score 32; DB 3; Length 152;  
Best Local Similarity 60.0%; Pred. No. 44;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMHY 10  
||| | | |  
Db 20 EEVFPPLAMNY 29

RESULT 11  
US-07-667-711B-4  
; Sequence 4, Application US/07667711B  
; Patent No. 6110700  
; GENERAL INFORMATION:  
; APPLICANT: ARNOLD, ANDREW  
; TITLE OF INVENTION: Pradl Cyclin and Its cDNA  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
; STREET: 1100 NEW YORK AVE., NW, SUITE 600  
; CITY: WASHINGTON  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/667,711B  
; FILING DATE: 11-MAR-1991  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: MCPHAIL, DONALD R.  
; REGISTRATION NUMBER: 35,811  
; REFERENCE/DOCKET NUMBER: 0609.4070000  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 371-2500  
; TELEFAX: (202) 371-2540  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 152 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: not relevant  
; MOLECULE TYPE: peptide  
US-07-667-711B-4

Query Match 57.1%; Score 32; DB 3; Length 152;  
Best Local Similarity 60.0%; Pred. No. 44;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMHY 10  
||| | | |  
Db 20 EEVFPPLAMNY 29

RESULT 12  
US-08-193-977-7  
; Sequence 7, Application US/08193977  
; Patent No. 5625031  
; GENERAL INFORMATION:  
; APPLICANT: WEBSTER, KEVIN R.  
; APPLICANT: COLEMAN, KEVIN G.  
; TITLE OF INVENTION: PEPTIDE INHIBITORS OF THE P33CDK2 AND  
; TITLE OF INVENTION: P34CDK2 CELL CYCLE REGULATORY KINASES AND HUMAN  
; TITLE OF INVENTION: PAPILLOMAVIRUS E7 ONCOPROTEIN

NUMBER OF SEQUENCES: 34  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: REED & ROBINS  
STREET: 635 BRYANT STREET  
CITY: PALO ALTO  
STATE: CALIFORNIA  
COUNTRY: UNITED STATES OF AMERICA  
ZIP: 94301  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/193,977  
FILING DATE: 08-FEB-1994  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: ROBINS, ROBERTA L.  
REGISTRATION NUMBER: 33,208  
REFERENCE/DOCKET NUMBER: 5998-0016  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 617-8999  
TELEFAX: (415) 327-3231  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 173 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-193-977-7

Query Match 57.1%; Score 32; DB 1; Length 173;  
Best Local Similarity 60.0%; Pred. No. 51;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMHY 10  
||| | | |  
Db 55 EEVFPPLAMNY 64

RESULT 13  
US-08-464-517-21  
; Sequence 21, Application US/08464517  
; Patent No. 5865640  
; GENERAL INFORMATION:  
; APPLICANT: BEACH, David H.  
; TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO  
; NUMBER OF SEQUENCES: 50  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD  
; STREET: 60 State Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII(text)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/464,517  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/963,308  
FILING DATE: 16-OCT-1992  
APPLICATION NUMBER: US 07/888,178  
FILING DATE: 26-MAY-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/701,514

;; FILING DATE: 16-MAY-1991  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Matthew P. Vincent  
;; REGISTRATION NUMBER: 36,709  
;; REFERENCE/DOCKET NUMBER: MII-004C  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (617) 227-7400  
;; TELEFAX: (617) 227-5941  
;; INFORMATION FOR SEQ ID NO: 21:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 189 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
US-08-464-517-21

Query Match 57.1%; Score 32; DB 2; Length 189;  
Best Local Similarity 60.0%; Pred. No. 56;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
QY 1 EEVVPXGMHY 10  
Db 74 EEVFLPMNY 83

RESULT 14  
US-08-246-361A-21  
; Sequence 21, Application US/08246361A  
; Patent No. 5998582  
; GENERAL INFORMATION:  
; APPLICANT: BEACH, David H.  
; TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO  
; NUMBER OF SEQUENCES: 50  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD  
; STREET: 60 State Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII(text)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/246,361A  
; FILING DATE: 19-MAY-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/963,308  
; FILING DATE: 16-OCT-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/888,178  
; FILING DATE: 26-MAY-1992  
; APPLICATION NUMBER: US 07/701,514  
; FILING DATE: 16-MAY-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Matthew P. Vincent  
; REGISTRATION NUMBER: 36,709  
; REFERENCE/DOCKET NUMBER: MII-004C  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 227-7400  
; TELEFAX: (617) 227-5941  
; INFORMATION FOR SEQ ID NO: 21:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 189 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide

US-08-246-361A-21

Query Match 57.1%; Score 32; DB 2; Length 189;  
Best Local Similarity 60.0%; Pred. No. 56;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
QY 1 EEVVPXGMHY 10  
Db 74 EEVFLPMNY 83

RESULT 15  
US-08-463-772-21  
; Sequence 21, Application US/08463772  
; Patent No. 6066501  
; GENERAL INFORMATION:  
; APPLICANT: BEACH, David H.  
; TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO  
; NUMBER OF SEQUENCES: 50  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD  
; STREET: 60 State Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII(text)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/463,772  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/963,308  
; FILING DATE: 16-OCT-1992  
; APPLICATION NUMBER: US 07/888,178  
; FILING DATE: 26-MAY-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/701,514  
; FILING DATE: 16-MAY-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Matthew P. Vincent  
; REGISTRATION NUMBER: 36,709  
; REFERENCE/DOCKET NUMBER: MII-004C  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 227-7400  
; TELEFAX: (617) 227-5941  
; INFORMATION FOR SEQ ID NO: 21:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 189 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-463-772-21

Query Match 57.1%; Score 32; DB 3; Length 189;  
Best Local Similarity 60.0%; Pred. No. 56;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
QY 1 EEVVPXGMHY 10  
Db 74 EEVFLPMNY 83

RESULT 16  
PCT-US93-05000-21  
; Sequence 21, Application PC/TUS9305000  
; GENERAL INFORMATION:  
; APPLICANT: MITOTIX

;; TITLE OF INVENTION: D-Type Cyclin and Uses Related Thereto  
;; NUMBER OF SEQUENCES: 42  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
;; STREET: Two Militia Drive  
;; CITY: Lexington  
;; STATE: Massachusetts  
;; COUNTRY: US  
;; ZIP: 02173  
;;  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: Patent Release #1.0, Version #1.25  
;;  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: PCT/US93/05000  
;; FILING DATE: 19930525  
;; CLASSIFICATION:  
;;  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US/07/888,178  
;; FILING DATE: 26-MAY-1992  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Granahan, Patricia  
;; REGISTRATION NUMBER: 32,227  
;; REFERENCE/DOCKET NUMBER: CSHL91-02A  
;; TELEPHONE: 617-861-6240  
;; TELEFAX: 616-861-9540  
;;  
;; INFORMATION FOR SEQ ID NO: 21:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 189 amino acids  
;; TYPE: AMINO ACID  
;; TOPOLOGY: unknown  
;; MOLECULE TYPE: protein  
;; PCT-US93-05000-21

Query Match 57.1%; Score 32; DB 5; Length 189;  
Best Local Similarity 60.0%; Pred. No. 56;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMHY 10  
||| | | |  
Db 74 EEVFPPLAMNY 83

RESULT 17  
US-08-464-517-22  
; Sequence 22, Application US/08464517  
; Patent No. 5869640  
; GENERAL INFORMATION:  
; APPLICANT: BEACH, David H.  
; TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO  
; NUMBER OF SEQUENCES: 50  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD  
; STREET: 60 State Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII(text)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/464,517  
; FILING DATE: 19-MAY-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/963,308  
; FILING DATE: 16-OCT-1992  
; TELEPHONE: (617) 227-7400  
; TELEFAX: (617) 227-5941  
; NAME: Matthew P. Vincent  
; REGISTRATION NUMBER: 36,709  
; REFERENCE/DOCKET NUMBER: MII-004C  
; TELECOMMUNICATION INFORMATION:  
; INFORMATION FOR SEQ ID NO: 22:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 236 amino acids  
; TYPE: amino acid

;; FILING DATE: 26-MAY-1992  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/701,514  
;; FILING DATE: 16-MAY-1991  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Matthew P. Vincent  
;; REGISTRATION NUMBER: 36,709  
;; REFERENCE/DOCKET NUMBER: MII-004C  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (617) 227-7400  
;; TELEFAX: (617) 227-5941  
;; INFORMATION FOR SEQ ID NO: 22:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 236 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
;; US-08-464-517-22

Query Match 57.1%; Score 32; DB 2; Length 236;  
Best Local Similarity 60.0%; Pred. No. 72;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMHY 10  
||| | | |  
Db 20 EEVFPPLAMNY 29

RESULT 18  
US-08-246-361A-22  
; Sequence 22, Application US/08246361A  
; Patent No. 5998582  
; GENERAL INFORMATION:  
; APPLICANT: BEACH, David H.  
; TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO  
; NUMBER OF SEQUENCES: 50  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD  
; STREET: 60 State Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII(text)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/246,361A  
; FILING DATE: 19-MAY-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/963,308  
; FILING DATE: 16-OCT-1992  
; TELEPHONE: (617) 227-7400  
; TELEFAX: (617) 227-5941  
; NAME: Matthew P. Vincent  
; REGISTRATION NUMBER: 36,709  
; REFERENCE/DOCKET NUMBER: MII-004C  
; TELECOMMUNICATION INFORMATION:  
; INFORMATION FOR SEQ ID NO: 22:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 236 amino acids  
; TYPE: amino acid

```

; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-246-361A-22

```

```

Query Match          57.1%; Score 32; DB 2; Length 236;
Best Local Similarity 60.0%; Pred. No. 72;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

```

```

QY      1 EEVVPXGMHY 10
Db      20 EEVFLPMNY 29

```

## RESULT 19

```

US-08-463-772-22
; Sequence 22, Application US/08463772
; Patent No. 6066501
; GENERAL INFORMATION:
; APPLICANT: BEACH, David H.
; TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109

```

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII(text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/463,772
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/963,308
; FILING DATE: 16-OCT-1992
; APPLICATION NUMBER: US 07/888,178
; FILING DATE: 26-MAY-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/701,514
; FILING DATE: 16-MAY-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Matthew P. Vincent
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: MII-004C
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 236 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-463-772-22

```

```

Query Match          57.1%; Score 32; DB 3; Length 236;
Best Local Similarity 60.0%; Pred. No. 72;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

```

```

QY      1 EEVVPXGMHY 10
Db      20 EEVFLPMNY 29

```

## RESULT 20

```

PCT-US93-05000-22

```

```

; Sequence 22, Application PC/TUS9305000
; GENERAL INFORMATION:
; APPLICANT: MITORIX
; TITLE OF INVENTION: D-Type Cyclin and Uses Related Thereto
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: US
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/05000
; FILING DATE: 19930525
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/888,178
; FILING DATE: 26-MAY-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: CSHL91-02A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-6240
; TELEFAX: 616-861-9540
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 236 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
PCT-US93-05000-22

```

```

Query Match          57.1%; Score 32; DB 5; Length 236;
Best Local Similarity 60.0%; Pred. No. 72;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

```

```

QY      1 EEVVPXGMHY 10
Db      20 EEVFLPMNY 29

```

## RESULT 21

```

US-08-464-517-6
; Sequence 6, Application US/08464517
; Patent No. 5869640
; GENERAL INFORMATION:
; APPLICANT: BEACH, David H.
; TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII(text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/464,517
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:

```

APPLICATION NUMBER: US 07/963,308  
FILING DATE: 16-OCT-1992  
APPLICATION NUMBER: US 07/888,178  
FILING DATE: 26-MAY-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/701,514  
FILING DATE: 16-MAY-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Matthew P. Vincent  
REGISTRATION NUMBER: 36,709  
REFERENCE/DOCKET NUMBER: MII-004C  
TELEPHONE: (617) 227-7400  
TELEFAX: (617) 227-5941  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 280 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-464-517-6

Query Match 57.1%; Score 32; DB 2; Length 280;  
Best Local Similarity 60.0%; Pred. No. 87;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMHY 10  
||| | :|  
Db 75 EEVFPPLAMNY 84

RESULT 22  
US-08-463-772-6  
Sequence 6, Application US/08463772  
Patent No. 6066501  
GENERAL INFORMATION:  
APPLICANT: BEACH, David H.  
TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO  
NUMBER OF SEQUENCES: 50  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD  
STREET: 60 State Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII(text)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/463.772  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/963,308  
FILING DATE: 16-OCT-1992  
APPLICATION NUMBER: US 07/888,178  
FILING DATE: 26-MAY-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/701,514  
FILING DATE: 16-MAY-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Matthew P. Vincent  
REGISTRATION NUMBER: 36,709  
REFERENCE/DOCKET NUMBER: MII-004C  
TELEPHONE: (617) 227-7400  
TELEFAX: (617) 227-5941  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 280 amino acids

TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-463-772-6

Query Match 57.1%; Score 32; DB 3; Length 280;  
Best Local Similarity 60.0%; Pred. No. 87;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMHY 10  
||| | :|  
Db 75 EEVFPPLAMNY 84

RESULT 23  
US-08-246-361A-4  
Sequence 4, Application US/08246361A  
Patent No. 5998582  
GENERAL INFORMATION:  
APPLICANT: BEACH, David H.  
TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO  
NUMBER OF SEQUENCES: 50  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD  
STREET: 60 State Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII(text)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/246.361A  
FILING DATE: 19-MAY-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/963,308  
FILING DATE: 16-OCT-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/888,178  
FILING DATE: 26-MAY-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/701,514  
FILING DATE: 16-MAY-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Matthew P. Vincent  
REGISTRATION NUMBER: 36,709  
REFERENCE/DOCKET NUMBER: MII-004C  
TELEPHONE: (617) 227-7400  
TELEFAX: (617) 227-5941  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 289 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-246-361A-4

Query Match 57.1%; Score 32; DB 2; Length 289;  
Best Local Similarity 60.0%; Pred. No. 90;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMHY 10  
||| | :|  
Db 74 EEVFPPLAMNY 83

RESULT 24  
PCT-US93-05000-4

```
; Sequence 4, Application PC/TUS9305000
; GENERAL INFORMATION:
; APPLICANT: MITOTIX
; TITLE OF INVENTION: D-Type Cyclin and Uses Related Thereto
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: US
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/05000
; FILING DATE: 19930525
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/888,178
; FILING DATE: 26-MAY-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: CSHL91-02A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-9540
; TELEFAX: 616-861-9540
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 289 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; PCT-US93-05000-4
```

```
Query Match 57.1%; Score 32; DB 5; Length 289;
Best Local Similarity 60.0%; Pred. No. 90;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMHY 10
Db 74 EEVFPPLANNY 83
||| | | | |
```

```
RESULT 25
PCT-US93-05000-6
; Sequence 6, Application PC/TUS9305000
; GENERAL INFORMATION:
; APPLICANT: MITOTIX
; TITLE OF INVENTION: D-Type Cyclin and Uses Related Thereto
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: US
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/05000
; FILING DATE: 19930525
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/888,178
```

```
; FILING DATE: 26-MAY-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: CSHL91-02A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-6240
; TELEFAX: 616-861-9540
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 291 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; PCT-US93-05000-6
```

```
Query Match 57.1%; Score 32; DB 5; Length 291;
Best Local Similarity 60.0%; Pred. No. 91;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMHY 10
Db 75 EEVFPPLANNY 84
||| | | | |
```

```
Search completed: June 10, 2003, 13:51:31
Job time : 9.64286 secs
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2003, 13:46:50 ; Search time 15 Seconds

(without alignments)  
75.710 Million cell updates/sec

Title: US-09-909-164-7  
Perfect score: 56  
Sequence: 1 BEVVPXGMHYS 11

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 392085 seqs, 103240269 residues

Total number of hits satisfying chosen parameters: 392085

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 75 summaries

Database : Published Applications AA.\*

1: /cgn2.6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
2: /cgn2.6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2.6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2.6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2.6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2.6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
7: /cgn2.6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*  
8: /cgn2.6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2.6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*  
10: /cgn2.6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*  
11: /cgn2.6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*  
12: /cgn2.6/ptodata/2/pubpaa/US10\_PUBCOMB.pep.\*  
13: /cgn2.6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*  
14: /cgn2.6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	33	58.9	567	9	US-10-270-333-126
2	33	58.9	622	9	US-09-738-626-4919
3	33	58.9	747	9	US-09-866-050A-663
4	33	58.9	3472	9	US-10-027-806-4
5	33	58.9	3472	9	US-10-034-623-4
6	33	58.9	3472	9	US-10-027-801-4
7	32	57.1	254	10	US-09-778-927A-53
8	32	57.1	289	9	US-10-024-066-2
9	32	57.1	289	9	US-10-024-066-4
10	32	57.1	289	10	US-09-919-497-54
11	32	57.1	653	9	US-09-820-843A-26
12	32	57.1	715	9	US-09-252-088-16
13	32	57.1	793	9	US-09-252-088-15
14	31.5	56.2	662	9	US-10-047-542-80
15	31.5	56.2	847	9	US-09-870-759-52
16	31	55.4	59	10	US-09-948-080-14
17	31	55.4	773	10	US-09-864-761-40832
18	31	55.4	192	9	US-10-001-857-119
19	31	55.4	192	9	US-09-986-480-171

20	31	55.4	260	10	US-09-815-242-13489	Sequence 13489, A
21	31	55.4	260	10	US-09-815-242-13613	Sequence 13613, A
22	31	55.4	299	10	US-09-815-242-10697	Sequence 10697, A
23	31	55.4	336	9	US-09-782-974C-86	Sequence 86, Appl
24	31	55.4	337	9	US-10-023-775B-2	Sequence 2, Appl
25	31	55.4	337	9	US-10-270-144-2	Sequence 2, Appl
26	31	55.4	337	9	US-10-188-405-8	Sequence 8, Appl
27	31	55.4	337	9	US-09-885-453-1	Sequence 1, Appl
28	31	55.4	337	10	US-09-943-798-4	Sequence 4, Appl
29	31	55.4	447	10	US-09-765-272-182	Sequence 182, App
30	31	55.4	449	9	US-09-738-626-5315	Sequence 315, App
31	31	55.4	484	9	US-09-769-787-38	Sequence 38, Appl
32	31	55.4	529	10	US-09-923-304-4	Sequence 4, Appl
33	31	55.4	763	10	US-09-765-272-66	Sequence 66, Appl
34	31	55.4	796	10	US-09-765-272-56	Sequence 56, Appl
35	31	55.4	826	9	US-09-769-787-194	Sequence 194, App
36	31	55.4	838	9	US-09-884-465A-8	Sequence 8, Appl
37	31	55.4	840	9	US-09-884-465A-7	Sequence 7, Appl
38	31	55.4	840	9	US-09-884-465A-6	Sequence 10, Appl
39	31	55.4	1039	9	US-09-884-465A-10	Sequence 6, Appl
40	31	55.4	1463	9	US-10-176-847-22	Sequence 22, Appl
41	31	55.4	1724	9	US-09-964-899-43	Sequence 43, Appl
42	30	53.6	7	9	US-09-909-062-1	Sequence 1, Appl
43	30	53.6	7	9	US-09-909-062-9	Sequence 9, Appl
44	30	53.6	7	9	US-09-909-062-130	Sequence 130, App
45	30	53.6	44	10	US-09-881-752A-292	Sequence 292, App
46	30	53.6	72	10	US-09-925-297-664	Sequence 664, App
47	30	53.6	121	9	US-09-852-797-68	Sequence 68, Appl
48	30	53.6	121	9	US-09-852-797-85	Sequence 85, Appl
49	30	53.6	121	10	US-09-853-161-68	Sequence 88, Appl
50	30	53.6	121	10	US-09-853-161-85	Sequence 85, Appl
51	30	53.6	121	10	US-09-852-659A-68	Sequence 68, Appl
52	30	53.6	121	10	US-09-852-659A-85	Sequence 85, Appl
53	30	53.6	135	9	US-09-932-598-359	Sequence 359, App
54	30	53.6	135	9	US-09-989-293A-359	Sequence 359, App
55	30	53.6	135	9	US-09-989-735-359	Sequence 359, App
56	30	53.6	135	9	US-09-990-444-359	Sequence 359, App
57	30	53.6	135	9	US-09-989-730-359	Sequence 359, App
58	30	53.6	135	9	US-09-990-436-359	Sequence 359, App
59	30	53.6	135	9	US-09-991-181-359	Sequence 359, App
60	30	53.6	135	9	US-09-993-687-359	Sequence 359, App
61	30	53.6	135	9	US-09-989-734-359	Sequence 359, App
62	30	53.6	135	9	US-09-997-653-359	Sequence 359, App
63	30	53.6	135	9	US-10-174-590-444	Sequence 444, App
64	30	53.6	135	9	US-10-176-758-444	Sequence 444, App
65	30	53.6	135	9	US-10-175-737-444	Sequence 444, App
66	30	53.6	135	9	US-09-993-667-359	Sequence 359, App
67	30	53.6	135	9	US-10-173-706-444	Sequence 444, App
68	30	53.6	135	9	US-10-175-738-444	Sequence 444, App
69	30	53.6	135	9	US-10-175-752-444	Sequence 444, App
70	30	53.6	135	9	US-10-176-482-444	Sequence 444, App
71	30	53.6	135	9	US-10-176-757-444	Sequence 444, App
72	30	53.6	135	9	US-10-176-913-444	Sequence 444, App
73	30	53.6	135	9	US-10-180-552-444	Sequence 444, App
74	30	53.6	135	9	US-10-180-557-444	Sequence 444, App
75	30	53.6	135	9	US-09-990-438-359	Sequence 359, App

## ALIGNMENTS

### RESULT 1

US-10-270-333-126  
; Sequence 126, Application US/10270333  
; Publication No. US20030092124A1  
; GENERAL INFORMATION:  
; APPLICANT: Cravchik, Anibal  
; TITLE OF INVENTION: ISOLATED G-PROTEIN COUPLED RECEPTORS,  
NUCLEIC ACID MOLECULES ENCODING GPCR PROTEINS, AND USES  
; TITLE OF INVENTION: THEROF AS INSECTICIDAL TARGETS  
; FILE REFERENCE: CLO00733CON  
; CURRENT APPLICATION NUMBER: US/10/270,333  
; CURRENT FILING DATE: 2002-10-15

; PRIOR APPLICATION NUMBER: 60/168,677  
 ; PRIOR FILING DATE: 1999-12-03  
 ; PRIOR APPLICATION NUMBER: 60/175,691  
 ; PRIOR FILING DATE: 2000-01-12  
 ; PRIOR APPLICATION NUMBER: 60/191,638  
 ; PRIOR FILING DATE: 2000-03-23  
 ; NUMBER OF SEQ ID NOS: 198  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 126  
 ; LENGTH: 567  
 ; TYPE: PRT  
 ; ORGANISM: Drosophila  
 ; US-10-270-333-126

Query Match 58.9%; Score 33; DB 9; Length 567;  
 Best Local Similarity 83.3%; Pred. No. 2.8e+02;  
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 PXGMHY 10  
 Db 402 PSGMHY 407

RESULT 2  
 US-09-738-626-4919  
 ; Sequence 4919, Application US/09738626  
 ; Publication No. US20020197605A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: NAKAGAWA, SATOSHI  
 ; APPLICANT: MIZOGUCHI, HIROSHI  
 ; APPLICANT: ANDO, SEIRO  
 ; APPLICANT: HAYASHI, MIKIRO  
 ; APPLICANT: OCHIAI, KEIKO  
 ; APPLICANT: YOKOI, HARUHIKO  
 ; APPLICANT: TATEISHI, NAOKO  
 ; APPLICANT: SENOH, AKIHIRO  
 ; APPLICANT: IKEDA, MASATO  
 ; APPLICANT: OZAKI, AKIO  
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
 ; FILE REFERENCE: 249-125  
 ; CURRENT APPLICATION NUMBER: US/09/738,626  
 ; CURRENT FILING DATE: 2000-12-18  
 ; PRIOR APPLICATION NUMBER: JP 99/377484  
 ; PRIOR FILING DATE: 1999-12-16  
 ; PRIOR APPLICATION NUMBER: JP 00/159162  
 ; PRIOR FILING DATE: 2000-04-07  
 ; PRIOR APPLICATION NUMBER: JP 00/280988  
 ; NUMBER OF SEQ ID NOS: 2000-08-03  
 ; SOFTWARE: PatentIn ver. 3.0  
 ; SEQ ID NO 4919  
 ; LENGTH: 622  
 ; TYPE: PRT  
 ; ORGANISM: Corynebacterium glutamicum  
 ; US-09-738-626-4919

Query Match 58.9%; Score 33; DB 9; Length 622;  
 Best Local Similarity 50.0%; Pred. No. 3.1e+02;  
 Matches 6; Conservative 3; Mismatches 1; Indels 2; Gaps 1;

QY 1 EEVVPXGM--HY 10  
 Db 46 EEIPIGVPNHY 57

RESULT 3  
 US-09-866-050A-663  
 ; Sequence 663, Application US/09866050A  
 ; Publication No. US20030040471A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Watson, James D.  
 ; APPLICANT: Strachan, Lorna  
 ; APPLICANT: Sleeman, Matthew

; APPLICANT: Onrust, Rene  
 ; APPLICANT: Murison, James G.  
 ; APPLICANT: Kumble, Krishanand D.  
 ; TITLE OF INVENTION: Compositions Isolated From Skin Cells  
 ; FILE REFERENCE: 11000.1011c4U  
 ; CURRENT APPLICATION NUMBER: US/09/866,050A  
 ; CURRENT FILING DATE: 2001-05-24  
 ; NUMBER OF SEQ ID NOS: 725  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 663  
 ; LENGTH: 747  
 ; TYPE: PRT  
 ; ORGANISM: Rat  
 ; US-09-866-050A-663

Query Match 58.9%; Score 33; DB 9; Length 747;  
 Best Local Similarity 71.4%; Pred. No. 3.8e+02;  
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 PXGMHYS 11  
 Db 627 PGLHYS 633

RESULT 4  
 US-10-027-806-4  
 ; Sequence 4, Application US/10027806  
 ; Patent No. US20020160476A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Swanson, Ronald V.  
 ; APPLICANT: Feldman, Robert A.  
 ; APPLICANT: Schleper, Christa  
 ; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM CENARCHAEUM SYMBIOSUM  
 ; FILE REFERENCE: DCRP.002A  
 ; CURRENT APPLICATION NUMBER: US/10/027,806  
 ; CURRENT FILING DATE: 2001-12-21  
 ; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/408,020  
 ; PRIOR FILING DATE: EARLIER FILING DATE: 1999-09-29  
 ; NUMBER OF SEQ ID NOS: 123  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 4  
 ; LENGTH: 3472  
 ; TYPE: PRT  
 ; ORGANISM: Cenarchaeum symbiosum  
 ; US-10-027-806-4

Query Match 58.9%; Score 33; DB 9; Length 3472;  
 Best Local Similarity 45.5%; Pred. No. 2.1e+03;  
 Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMHYS 11  
 Db 2294 EDVIPRGISFS 2304

RESULT 5  
 US-10-034-623-4  
 ; Sequence 4, Application US/10034623  
 ; Publication No. US20020198365A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Swanson, Ronald V.  
 ; APPLICANT: Feldman, Robert A.  
 ; APPLICANT: Schleper, Christa  
 ; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM CENARCHAEUM SYMBIOSUM  
 ; FILE REFERENCE: DCRP.002A  
 ; CURRENT APPLICATION NUMBER: US/10/034,623  
 ; CURRENT FILING DATE: 2001-12-21  
 ; PRIOR APPLICATION NUMBER: 09/408,020  
 ; PRIOR FILING DATE: 1999-09-29  
 ; PRIOR APPLICATION NUMBER: 60/102,294  
 ; PRIOR FILING DATE: 1998-09-29  
 ; NUMBER OF SEQ ID NOS: 123

: SOFTWARE: FastSEO for Windows Version 3.0

```

; SOFTWARE: FASTSEQ FOR WINDOWS V
; SEQ ID NO 4
; LENGTH: 3472
; TYPE: PRT
; ORGANISM: Cenarchaeum symbiosum
US-10-034-623-4

```

Query Match 58.9%; Score 33; DB 9; Length 3472;  
Best Local Similarity 45.5%; Pred. No. 2.1e+03;  
Matches 5; Conservative 4; Mismatches 2; Indels

```

QY      1 EEVVPXGMHYS 11
        | : | : | : |
Db     2294 EDVIPRGISES 2304

```

RESULT 6  
US-10-027-801-4 ; Sequence 4, Application US/10027801  
; Publication NO. US20030054364A1  
: GENERAL INFORMATION:

Query Match 58.9%; Score 33; DB 9; Length 3472;  
Best Local Similarity 45.5%; Pred. No. 2.1e+03;  
Matches 5; Conservative 4; Mismatches 2; Indels

```
QY      1 EEVVPXGMHYS 11
        | : | | : |
nb     2294 EDVTPRGISES 2304
```

RESULT 7  
US-09-778-927A-53  
; Sequence 53, Application US/09778927A  
; Patent No. US20020068342A1  
; GENERAL INFORMATION:

```
Query Match 57.1%; Score 32; DB 10; Length 254;
Best Local Similarity 60.0%; pred. No. 1.8e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
```

QY 1 EEVVPXGMHY 10  
||| | | : |  
74 EEVEPIAMNY 83

```

RESULT 8
US-10-024-066-2
; Sequence 2, Application US/10024066
; Patent No. US20020166134A1
; GENERAL INFORMATION:
; APPLICANT: Field, Loren J.
; APPLICANT: Pasumarthi, Kishore Babu S.
; TITLE OF INVENTION: CARDIOMYOCYTES WITH ENHANCED PROLIFERATIVE POTENTIAL,
; TITLE OF INVENTION: AND METHODS FOR PREPARING AND USING SAME
; FILE REFERENCE: 7037-450
; CURRENT APPLICATION NUMBER: US/10/024,066
; CURRENT FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: 60/139,942
; PRIOR FILING DATE: 1999-06-18
; PRIOR APPLICATION NUMBER: PCT/US00/16927
; PRIOR FILING DATE: 2000-06-19
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 289
; TYPE: prt
; ORGANISM: Mus musculus
; US-10-024-066-2

```

Query Match 57.1%; Score 32; DB 9; Length 289;  
Best Local Similarity 60.0%; Pred. No. 2.1e+02;  
Matches 6: Conservative 1; Mismatches 3; Indels

QY 1 EEVVPXGMHY 10  
||| | | : |  
Db 73 EEVFPAMNY 82

RESULT 9  
US-10-024-066-4  
; Sequence 4, Application US/10024066  
; Patent NO. US20020166134A1  
; GENERAL INFORMATION:  
; APPLICANT: Field, Loren J.  
; APPLICANT: Pasumarthi, Kishore Babu S.  
; TITLE OF INVENTION: CARDIOMYOCYTES WITH ENHANCED PROLIFERATIVE POTENTIAL,  
; TITLE OF INVENTION: AND METHODS FOR PREPARING AND USING SAME  
; FILE REFERENCE: 7037-450  
; CURRENT APPLICATION NUMBER: US/10/024,066  
; CURRENT FILING DATE: 2001-12-18  
; PRIOR APPLICATION NUMBER: 60/139,942  
; PRIOR FILING DATE: 1999-06-18  
; PRIOR APPLICATION NUMBER: PCT/US00/16827  
; PRIOR FILING DATE: 2000-06-19  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 289  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-024-066-4

Query Match 57.1%; Score 32; DB 9; Length 289;  
Best Local Similarity 60.0%; Pred. No. 2.1e+02;  
Matches 6; Conservative 1; Mismatches 3; Indels

QY 1 EEVVPXGMHY 10  
||| | |

Db 74 EEVFLPMNY 83

RESULT 10

US-09-919-497-54  
; Sequence 54, Application US/09919497  
; Patent No. US20020106662A1  
; GENERAL INFORMATION:  
; APPLICANT: Mutter, George L.  
; TITLE OF INVENTION: PROGNOSTIC CLASSIFICATION OF ENDOMETRIAL CANCER  
; FILE REFERENCE: B0801/7225  
; CURRENT APPLICATION NUMBER: US/09/919,497  
; PRIOR FILING DATE: 2001-07-31  
; PRIOR APPLICATION NUMBER: US 60/221,735  
; PRIOR FILING DATE: 2000-07-31  
; NUMBER OF SEQ ID NOS: 100  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 54  
; LENGTH: 289  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-919-497-54

Query Match 57.1%; Score 32; DB 10; Length 289;  
Best Local Similarity 60.0%; Pred. No. 2.1e+02;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMHY 10

Db 74 EEVFLPMNY 83

RESULT 11

US-09-820-843A-26  
; Sequence 26, Application US/09820843A  
; Publication No. US20030039963A1  
; GENERAL INFORMATION:  
; APPLICANT: Council of Scientific and Industrial Research  
; TITLE OF INVENTION: A COMPUTATIONAL METHOD FOR THE IDENTIFICATION OF CANDIDATE PROTEIN  
; FILE REFERENCE: Q63915  
; CURRENT APPLICATION NUMBER: US/09/820,843A  
; CURRENT FILING DATE: 2001-03-30  
; NUMBER OF SEQ ID NOS: 118  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 26  
; LENGTH: 653  
; TYPE: PRT  
; ORGANISM: Vibrio cholerae  
; NAME/KEY: misc\_feature  
; OTHER INFORMATION: Iron(III) ABC transporter, permease protein  
; NAME/KEY: misc\_feature  
; OTHER INFORMATION: g1|9654609  
US-09-820-843A-26

Query Match 57.1%; Score 32; DB 9; Length 653;  
Best Local Similarity 75.0%; Pred. No. 5.1e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXGM 8

Db 300 EEVPSGI 307

RESULT 12

US-09-252-088-16  
; Sequence 16, Application US/09252088  
; Publication No. US20030031682A1  
; GENERAL INFORMATION:  
; APPLICANT: BRODEUR, Bernard R.  
; APPLICANT: BOYER, Martine  
; TITLE OF INVENTION: NOVEL IMMUNOADHESINS FOR TREATING AND PREVENTING VIRAL  
; FILE REFERENCE: 030905.0004.C1P1  
; CURRENT APPLICATION NUMBER: US/10/047,542  
; CURRENT FILING DATE: 2001-10-26  
; PRIOR APPLICATION NUMBER: PCT/US01/13932

; APPLICANT: CHARLEBOIS, Isabelle  
; APPLICANT: HAMEL, Jose  
; APPLICANT: MARTIN, Denis  
; TITLE OF INVENTION: NOVEL GROUP B STREPTOCOCCUS ANTIGENS  
; FILE REFERENCE: 8331-9002  
; CURRENT APPLICATION NUMBER: US/09/252,088  
; CURRENT FILING DATE: 1999-02-18  
; EARLIER APPLICATION NUMBER: US/60/075,425  
; EARLIER FILING DATE: 1998-02-20  
; NUMBER OF SEQ ID NOS: 44  
; SEQ ID NO 16  
; LENGTH: 715  
; TYPE: PRT  
; ORGANISM: group B streptococcus  
US-09-252-088-16

Query Match 57.1%; Score 32; DB 9; Length 715;  
Best Local Similarity 75.0%; Pred. No. 5.7e+02;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 VVPXGMHY 10

Db 243 VVPHGDHY 250

RESULT 13

US-09-252-088-15  
; Sequence 15, Application US/09252088  
; Publication No. US20030031682A1  
; GENERAL INFORMATION:  
; APPLICANT: BRODEUR, Bernard R.  
; APPLICANT: BOYER, Martine  
; APPLICANT: CHARLEBOIS, Isabelle  
; APPLICANT: HAMEL, Jose  
; APPLICANT: MARTIN, Denis  
; TITLE OF INVENTION: NOVEL GROUP B STREPTOCOCCUS ANTIGENS  
; FILE REFERENCE: 8331-9002  
; CURRENT APPLICATION NUMBER: US/09/252,088  
; CURRENT FILING DATE: 1999-02-18  
; EARLIER APPLICATION NUMBER: US/60/075,425  
; EARLIER FILING DATE: 1998-02-20  
; NUMBER OF SEQ ID NOS: 44  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 15  
; LENGTH: 793  
; TYPE: PRT  
; ORGANISM: group B streptococcus  
US-09-252-088-15

Query Match 57.1%; Score 32; DB 9; Length 793;  
Best Local Similarity 75.0%; Pred. No. 6.4e+02;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 VVPXGMHY 10

Db 321 VVPHGDHY 328

RESULT 14

US-10-047-542-80  
; Sequence 80, Application US/10047542  
; Patent No. US20020168367A1  
; GENERAL INFORMATION:  
; APPLICANT: LARRICK, JAMES W.  
; APPLICANT: WYCOFF, KEITH L.  
; TITLE OF INVENTION: NOVEL IMMUNOADHESINS FOR TREATING AND PREVENTING VIRAL  
; FILE REFERENCE: 030905.0004.C1P1  
; CURRENT APPLICATION NUMBER: US/10/047,542  
; CURRENT FILING DATE: 2001-10-26  
; PRIOR APPLICATION NUMBER: PCT/US01/13932

us-09-909-164-7.rapb

Wed Jun 11 15:44:30 2003

Query Match 55.4%; Score 31.5; DB 9; Length 59;  
Best Local Similarity 45.5%; Pred. No. 57;  
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;  
PQ 1 EEVVPXGMHYS 11  
DB 38 EKHIPGGLYS 48  
FILE REFERENCE: Aecomica-X-1  
CURRENT APPLICATION NUMBER: US/09/864,761  
CURRENT FILING DATE: 2001-05-23  
PRIOR APPLICATION NUMBER: US 60/180,312  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: US 60/207,456  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: US 09/632,366  
PRIOR FILING DATE: 2000-08-03  
PRIOR APPLICATION NUMBER: GB 24263.6  
PRIOR FILING DATE: 2000-10-04  
PRIOR APPLICATION NUMBER: US 60/236,359  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: PCT/US01/00666  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00667  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00664  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00669  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00665  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00668  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00663  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00662  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00661  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00670  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: US 60/234,687  
PRIOR FILING DATE: 2000-09-21  
PRIOR APPLICATION NUMBER: US 09/608,408  
PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: US 09/774,203  
PRIOR FILING DATE: 2001-01-29  
NUMBER OF SEQ ID NOS: 49117  
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
SEQ ID NO 40832  
LENGTH: 73  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO ALI36528.9  
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 4  
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.1  
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.5  
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3  
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.9  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.2

RESULT 17  
US-09-864-761-40832  
Sequence 40832, Application US/09864761  
Patent No. US20020048763A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharron G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
APPLICANT: Chen, Wensheng  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL  
TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY  
FILE REFERENCE: Aecomica-X-1  
CURRENT APPLICATION NUMBER: US/09/864,761  
CURRENT FILING DATE: 2001-05-23  
PRIOR APPLICATION NUMBER: US 60/180,312  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: US 60/207,456  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: US 09/632,366  
PRIOR FILING DATE: 2000-08-03  
PRIOR APPLICATION NUMBER: GB 24263.6  
PRIOR FILING DATE: 2000-10-04  
PRIOR APPLICATION NUMBER: US 60/236,359  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: PCT/US01/00666  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00667  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00664  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00669  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00665  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00668  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00663  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00662  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00661  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00670  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: US 60/234,687  
PRIOR FILING DATE: 2000-09-21  
PRIOR APPLICATION NUMBER: US 09/608,408  
PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: US 09/774,203  
PRIOR FILING DATE: 2001-01-29  
NUMBER OF SEQ ID NOS: 49117  
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
SEQ ID NO 40832  
LENGTH: 73  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO ALI36528.9  
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 4  
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.1  
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.5  
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3  
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.9  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.2

Query Match 56.2%; Score 31.5; DB 9; Length 62;  
Best Local Similarity 43.8%; Pred. No. 6.5e+02;  
Matches 7; Conservative 2; Mismatches 2; Indels 5; Gaps 1;  
PQ 1 EEVVP-----XGMHYS 11  
DB 623 ENVIPDFPEDEGIHYS 638  
FILE REFERENCE: Aecomica-X-1  
CURRENT APPLICATION NUMBER: US/09/870,759  
CURRENT FILING DATE: 2002-01-14  
PRIOR APPLICATION NUMBER: US 60/208,128  
PRIOR FILING DATE: 2000-05-30  
NUMBER OF SEQ ID NOS: 166  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 52  
LENGTH: 847  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-870-759-52

RESULT 15  
US-09-870-759-52  
Sequence 52, Application US/09870759  
Patent No. US20020177551A1  
GENERAL INFORMATION:  
APPLICANT: Terman, David S  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE  
FILE REFERENCE: 870759  
CURRENT APPLICATION NUMBER: US/09/870,759  
CURRENT FILING DATE: 2002-01-14  
PRIOR APPLICATION NUMBER: US 60/208,128  
PRIOR FILING DATE: 2000-05-30  
NUMBER OF SEQ ID NOS: 166  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 52  
LENGTH: 847  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-870-759-52

Query Match 56.2%; Score 31.5; DB 9; Length 847;  
Best Local Similarity 43.8%; Pred. No. 8.5e+02;  
Matches 7; Conservative 2; Mismatches 2; Indels 5; Gaps 1;  
PQ 1 EEVVP-----XGMHYS 11  
DB 808 ENVIPDFPEDEGIHYS 823  
FILE REFERENCE: Aecomica-X-1  
CURRENT APPLICATION NUMBER: US/09/948,080  
CURRENT FILING DATE: 2001-09-06  
PRIOR APPLICATION NUMBER: US/08/963,851  
PRIOR FILING DATE: 1997-11-04  
NUMBER OF SEQ ID NOS: 35  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 14  
LENGTH: 59  
TYPE: PRT  
ORGANISM: Enterococcus faecalis  
US-09-948-080-14

RESULT 16  
US-09-948-080-14  
Sequence 14, Application US/09948080  
Patent No. US20020102702A1  
GENERAL INFORMATION:  
APPLICANT: VAN DER OSTEN, CLAU  
APPLICANT: HALKIER, TORDEN  
APPLICANT: ANDERSEN, CARSTEN  
APPLICANT: BAUDITZ, PETER  
APPLICANT: HANSEN, PETER KAMP  
TITLE OF INVENTION: PROTEASE VARIANTS AND COMPOSITIONS  
FILE REFERENCE: 4946,200-US  
CURRENT APPLICATION NUMBER: US/09/948,080  
CURRENT FILING DATE: 2001-09-06  
PRIOR APPLICATION NUMBER: US/08/963,851  
PRIOR FILING DATE: 1997-11-04  
NUMBER OF SEQ ID NOS: 35  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 14  
LENGTH: 59  
TYPE: PRT  
ORGANISM: Enterococcus faecalis  
US-09-948-080-14

OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.2  
 OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.8  
 OTHER INFORMATION: EST\_HUMAN HIT: AW027739.1, EVALUATE 7.00e-14  
 US-09-864-761-40832

Query Match 55.4%; Score 31; DB 10; Length 73;  
 Best Local Similarity 71.4%; Pred. No. 72;  
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMH 9  
 ||| |:  
 Db 60 WPSGLH 66

## RESULT 18

US-10-001-857-119  
 ; Sequence 119, Application US/10001857  
 ; Publication No. US20020183500A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Macina, Roberto  
 ; APPLICANT: Recipon, Hervé  
 ; APPLICANT: Chen, Sel-yu  
 ; APPLICANT: Sun, Yongming  
 ; APPLICANT: Liu, Chenghua  
 ; TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Proteins  
 ; FILE REFERENCE: DEX-0273  
 ; CURRENT APPLICATION NUMBER: US/10/001,857  
 ; CURRENT FILING DATE: 2001-11-20  
 ; PRIOR APPLICATION NUMBER: 60/252,054  
 ; PRIOR FILING DATE: 2000-11-20  
 ; NUMBER OF SEQ ID NOS: 208  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 119  
 ; LENGTH: 192  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-10-001-857-119

Query Match 55.4%; Score 31; DB 9; Length 192;  
 Best Local Similarity 57.1%; Pred. No. 21e+02;  
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMH 9  
 :||| |:  
 Db 145 IIPKGMH 151

## RESULT 19

US-09-986-480-171  
 ; Sequence 171, Application US/09986480  
 ; Publication No. US20030027999A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Rosen et al.  
 ; TITLE OF INVENTION: 143 Human Secreted Proteins  
 ; FILE REFERENCE: PS500P1  
 ; CURRENT APPLICATION NUMBER: US/09/986,480  
 ; CURRENT FILING DATE: 2001-11-08  
 ; PRIOR APPLICATION NUMBER: PCT/US00/12788  
 ; PRIOR FILING DATE: 2000-05-11  
 ; PRIOR FILING DATE: 1999-05-13  
 ; NUMBER OF SEQ ID NOS: 456  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 171  
 ; LENGTH: 192  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-986-480-171

Query Match 55.4%; Score 31; DB 9; Length 192;  
 Best Local Similarity 50.0%; Pred. No. 2.1e+02;  
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 EVVPXGMHYS 11  
 | :| |:  
 Db 52 ENIPEGLNYS 61

## RESULT 20

US-09-815-242-13489  
 ; Sequence 13489, Application US/09815242  
 ; Patent No. US20020061569A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Haselbeck, Robert  
 ; APPLICANT: Ohlsen, Karl L.  
 ; APPLICANT: Zyskind, Judith W.  
 ; APPLICANT: Wall, Daniel  
 ; APPLICANT: Trawick, John D.  
 ; APPLICANT: Carr, Grant J.  
 ; APPLICANT: Yamamoto, Robert T.  
 ; APPLICANT: Xu, H. Howard  
 ; TITLE OF INVENTION: Identification of Essential Genes in Prokaryotes  
 ; FILE REFERENCE: ELITRA.011A  
 ; CURRENT APPLICATION NUMBER: US/09/815,242  
 ; CURRENT FILING DATE: 2001-03-21  
 ; PRIOR APPLICATION NUMBER: 60/191,078  
 ; PRIOR FILING DATE: 2000-03-21  
 ; PRIOR APPLICATION NUMBER: 60/206,848  
 ; PRIOR FILING DATE: 2000-05-23  
 ; PRIOR APPLICATION NUMBER: 60/207,727  
 ; PRIOR FILING DATE: 2000-05-26  
 ; PRIOR APPLICATION NUMBER: 60/242,578  
 ; PRIOR FILING DATE: 2000-10-23  
 ; PRIOR APPLICATION NUMBER: 60/253,625  
 ; PRIOR FILING DATE: 2000-11-27  
 ; PRIOR APPLICATION NUMBER: 60/257,931  
 ; PRIOR FILING DATE: 2000-12-22  
 ; PRIOR APPLICATION NUMBER: 60/269,308  
 ; PRIOR FILING DATE: 2001-02-16  
 ; NUMBER OF SEQ ID NOS: 14110  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 13489  
 ; LENGTH: 260  
 ; TYPE: PRT  
 ; ORGANISM: Streptococcus pneumoniae  
 US-09-815-242-13489

Query Match 55.4%; Score 31; DB 10; Length 260;  
 Best Local Similarity 60.0%; Pred. No. 2.9e+02;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EVVPXGMHY 10  
 || | |:  
 Db 188 EEKVGEGVHY 197

## RESULT 21

US-09-815-242-13613  
 ; Sequence 13613, Application US/09815242  
 ; Patent No. US20020061569A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Haselbeck, Robert  
 ; APPLICANT: Ohlsen, Karl L.  
 ; APPLICANT: Zyskind, Judith W.  
 ; APPLICANT: Wall, Daniel  
 ; APPLICANT: Trawick, John D.  
 ; APPLICANT: Carr, Grant J.  
 ; APPLICANT: Yamamoto, Robert T.  
 ; APPLICANT: Xu, H. Howard  
 ; TITLE OF INVENTION: Identification of Essential Genes in Prokaryotes  
 ; FILE REFERENCE: ELITRA.011A  
 ; CURRENT APPLICATION NUMBER: US/09/815,242  
 ; CURRENT FILING DATE: 2001-03-21  
 ; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; NUMBER OF SEQ ID NOS: 14110  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 13613  
; LENGTH: 260  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
US-09-815-242-13613

Query Match 55.4%; Score 31; DB 10; Length 260;  
Best Local Similarity 60.0%; Pred. No. 2.9e+02;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMHY 10  
||| |  
DB 188 EEKVGEGVHY 197

RESULT 22  
US-09-815-242-10697  
; Sequence 10697, Application US/09815242  
; Patent No. US20020061569A1  
; GENERAL INFORMATION:  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari L.  
; APPLICANT: Zyskind, Judith W.  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John D.  
; APPLICANT: Carr, Grant J.  
; APPLICANT: Yamamoto, Robert T.  
; APPLICANT: Xu, H. Howard  
; TITLE OF INVENTION: Identification of Essential Genes in  
; FILE REFERENCE: ELITRA.011A  
; CURRENT APPLICATION NUMBER: US/09/815,242  
; CURRENT FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; NUMBER OF SEQ ID NOS: 14110  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 10697  
; LENGTH: 299  
; TYPE: PRT  
; ORGANISM: Enterococcus faecalis  
US-09-815-242-10697

Query Match 55.4%; Score 31; DB 10; Length 299;  
Best Local Similarity 40.0%; Pred. No. 3.4e+02;  
Matches 4; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMHY 10  
|:| |:  
DB 218 EOIITPTGIEY 227

RESULT 23  
US-09-782-974C-86  
; Sequence 86, Application US/09782974C  
; Publication No. US20030082534A1  
; GENERAL INFORMATION:  
; APPLICANT: Vogeli, Gabriel  
; APPLICANT: Lind, Peter  
; APPLICANT: Wood, Linda S.  
; APPLICANT: Parodi, Luis A.  
; TITLE OF INVENTION: No. US20030082534A1 G Protein Coupled Receptor  
; FILE REFERENCE: 411USPHRM311  
; CURRENT APPLICATION NUMBER: US/09/782,974C  
; CURRENT FILING DATE: 2002-09-04  
; PRIOR APPLICATION NUMBER: 60/165,838  
; PRIOR FILING DATE: 1999-11-16  
; PRIOR APPLICATION NUMBER: 09/714,449  
; PRIOR FILING DATE: 2000-11-16  
; PRIOR APPLICATION NUMBER: 60/198,568  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: 60/166,071  
; PRIOR FILING DATE: 1999-11-17  
; PRIOR APPLICATION NUMBER: 60/166,678  
; PRIOR FILING DATE: 1999-11-19  
; PRIOR APPLICATION NUMBER: 60/173,396  
; PRIOR FILING DATE: 1999-12-28  
; PRIOR APPLICATION NUMBER: 60/184,129  
; PRIOR FILING DATE: 2000-02-22  
; PRIOR APPLICATION NUMBER: 60/185,421  
; PRIOR FILING DATE: 2000-02-28  
; PRIOR APPLICATION NUMBER: 60/185,554  
; PRIOR FILING DATE: 2000-02-28  
; PRIOR APPLICATION NUMBER: 60/186,530  
; PRIOR FILING DATE: 2000-03-02  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 192  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 86  
; LENGTH: 336  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-782-974C-86

Query Match 55.4%; Score 31; DB 9; Length 336;  
Best Local Similarity 50.0%; Pred. No. 3.8e+02;  
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMHY 10  
|:| |:  
DB 26 DENIPLKMHY 35

RESULT 24  
US-10-023-775B-2  
; Sequence 2, Application US/10023775B  
; Publication No. US2003002282A1  
; GENERAL INFORMATION:  
; APPLICANT: Pfizer Ltd. (EP/GB) only  
; APPLICANT: Pfizer Inc. (US, JP, EB except GB)  
; APPLICANT: Fidock, Mark David  
; TITLE OF INVENTION: No. US2003002282A1 Polypeptide  
; FILE REFERENCE: PCI0959AGPR  
; CURRENT APPLICATION NUMBER: US/10/023,775B  
; CURRENT FILING DATE: 2001-12-18  
; PRIOR APPLICATION NUMBER: GB 0030854.4  
; PRIOR FILING DATE: 2000-12-18  
; PRIOR APPLICATION NUMBER: US 60/260,590  
; PRIOR FILING DATE: 2001-01-09

; PRIOR APPLICATION NUMBER: US 60/296,660  
; PRIOR FILING DATE: 2001-06-07  
; PRIOR APPLICATION NUMBER: GB 0111031.1  
; PRIOR FILING DATE: 2001-05-04  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2  
; LENGTH: 337  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-023-775B-2

Query Match 55.4%; Score 31; DB 9; Length 337;  
Best Local Similarity 50.0%; Pred. NO. 3.8e+02;  
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMHY 10  
Db :|:|  
26 DENIPLKMHY 35

## RESULT 25

US-10-270-144-2  
; Sequence 2, Application US/10270144  
; Publication No. US20030049790A1  
; GENERAL INFORMATION:  
; APPLICANT: WEI, Ming-Hui et al  
; TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED  
; TITLE OF INVENTION: RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR  
; TITLE OF INVENTION: PROTEINS, AND USES THEREOF  
; FILE REFERENCE: CL000750CON  
; CURRENT APPLICATION NUMBER: US/10/270,144  
; CURRENT FILING DATE: 2002-10-15  
; PRIOR APPLICATION NUMBER: 60/205,196  
; PRIOR FILING DATE: 2000-05-18  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 337  
; TYPE: PRT  
; ORGANISM: Human  
US-10-270-144-2

Query Match 55.4%; Score 31; DB 9; Length 337;  
Best Local Similarity 50.0%; Pred. NO. 3.8e+02;  
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMHY 10  
Db :|:|  
26 DENIPLKMHY 35

Search completed: June 10, 2003, 14:35:41  
Job time : 16.0714 secs



GenCore version 5.1.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2003, 13:31:15 ; Search time 11.2143 Seconds  
(without alignments)  
94.297 Million cell updates/sec

Title: US-09-909-164-7  
Perfect score: 56  
Sequence: 1 EEVVPXGMHYS 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 75 summaries

Database : PIR.73.\*

1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	37	66.1	308	2 A72207	fish proteinase ac
2	37	66.1	1057	2 F89892	carbamoyl-phosphat
3	36	64.3	102	2 A42452	VI protein - tobac
4	36	64.3	252	2 AE2001	hypothetical prote
5	36	64.3	460	2 S69046	hypothetical prote
6	36	64.3	743	2 S38143	hypothetical prote
7	35	62.5	156	2 D82618	conserved hypothet
8	35	62.5	233	2 T02590	DNA binding protei
9	35	62.5	311	2 H69194	GMP synthetase, su
10	35	62.5	425	2 T24111	hypothetical prote
11	35	62.5	510	2 G86430	T518.1 protein - A
12	34	60.7	264	2 G69117	diphthine synthase
13	34	60.7	279	2 C75538	hypothetical prote
14	34	60.7	350	2 B75478	3-dehydroquinatase
15	34	60.7	355	2 T35025	probable DNA ligas
16	34	60.7	360	2 E69086	cell division prot
17	34	60.7	425	2 C83903	hypothetical prote
18	34	60.7	426	2 S58132	sls1 protein precu
19	34	60.7	495	2 T28717	hypothetical prote
20	34	60.7	1028	2 AF3286	ATP-dependent DNA
21	33	58.9	156	2 S54619	hypothetical prote
22	33	58.9	367	2 E83607	polyamine transpor
23	33	58.9	441	2 G82253	conserved hypothet
24	33	58.9	466	2 G71542	probable amino aci
25	33	58.9	466	2 H81697	amino acid antipor
26	33	58.9	487	2 S65811	finger protein (cl
27	33	58.9	514	1 HQDVLB	cytochrome-c3 hydr
28	33	58.9	534	2 A69284	coenzyme F420-quin
29	33	58.9	545	2 T08564	hypothetical prote

#### ALIGNMENTS

##### RESULT 1

A72207  
fish proteinase activity modulator HflK - Thermotoga maritima (strain MSB8)  
C:Species: Thermotoga maritima  
C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 21-Jul-2000  
C:Accession: A72207  
R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; I  
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson  
C.M.

Nature 399, 323-329, 1999

A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genom

A:Reference number: A72200; MUID:99287316; PMID:10360571

A:Accession: A72207

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-308 <ARN>

A:Cross-references: GB:AE001819; GB:AE000512; NID:94982396; PIDN:AAD36885.1; PID:9

A:Experimental source: strain MSB8

C:Genetics:

A:Gene: TM1822

C:Superfamily: erythrocyte band 7 integral membrane protein

30	33	58.9	627	2	A69663
31	33	58.9	716	1	JC5061
32	33	58.9	1257	2	S44754
33	33	58.9	1396	2	S36851
34	33	58.9	3472	2	T31308
35	32	57.1	126	2	A25905
36	32	57.1	197	2	D71640
37	32	57.1	225	2	S57810
38	32	57.1	233	2	E97120
39	32	57.1	267	2	T07215
40	32	57.1	270	2	C95881
41	32	57.1	283	2	T25737
42	32	57.1	288	2	JC4011
43	32	57.1	288	2	I58372
44	32	57.1	289	2	A41984
45	32	57.1	289	2	A42822
46	32	57.1	291	2	S57922
47	32	57.1	291	2	S57922
48	32	57.1	291	2	JC4579
49	32	57.1	291	2	S62730
50	32	57.1	292	2	B42832
51	32	57.1	295	2	A38977
52	32	57.1	295	2	A56523
53	32	57.1	295	2	JC2342
54	32	57.1	335	2	T25498
55	32	57.1	341	2	S43354
56	32	57.1	354	2	JX0116
57	32	57.1	363	2	A83470
58	32	57.1	369	2	T17267
59	32	57.1	374	2	G69119
60	32	57.1	384	2	AD0049
61	32	57.1	405	2	E82626
62	32	57.1	412	2	A48702
63	32	57.1	509	2	S51348
64	32	57.1	519	2	G84598
65	32	57.1	534	2	F89263
66	32	57.1	617	2	AC1421
67	32	57.1	653	2	D82352
68	32	57.1	670	2	S22293
69	32	57.1	696	2	A91247
70	32	57.1	704	2	T29996
71	32	57.1	759	2	S25330
72	32	57.1	822	2	T46758
73	32	57.1	840	2	T39116
74	32	57.1	877	2	T40413
75	32	57.1	1369	2	T03104

Nature 399, 323-329, 1999

A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genom

A:Reference number: A72200; MUID:99287316; PMID:10360571

A:Accession: A72207

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-308 <ARN>

A:Cross-references: GB:AE001819; GB:AE000512; NID:94982396; PIDN:AAD36885.1; PID:9

A:Experimental source: strain MSB8

C:Genetics:

A:Gene: TM1822

C:Superfamily: erythrocyte band 7 integral membrane protein

DNA mismatch repair  
macrophage-stimula  
C14B9.8 protein -  
L-shaped tail fibe  
hypothetical 367K  
tachykinin B precu  
heme exporter prot  
hypothetical prote  
ribosomal protein  
probable Thua prot  
hypothetical prote  
cyclin D2 - rat  
cyclin D2 - mouse  
cyclin D2 - human  
cyclin D1 - Africa  
cyclin D2 - chicke  
cyclin D1 - zebra  
cyclin D3 - human  
cyclin D1 - human  
cyclin D1 - mouse  
cyclin D1 - rat  
hypothetical prote  
hypothetical prote  
type II site-speci  
probable periplasm  
hypothetical prote  
probable aspartate  
hypothetical prote  
hypothetical prote  
2-methyl-branched-  
hypothetical prote  
probable bZIP tran  
conserved hypothet  
beta-glucoside-spe  
iron(III) ABC tran  
zinc finger protei  
phage transposase  
hypothetical prote  
SCT1 protein - yea  
hypothetical 92.4K  
probable sulfate p  
sulfate permease -  
tegument protein h

Query Match 66.1%; Score 37; DB 2; Length 308;  
 Best Local Similarity 75.0%; Pred. No. 10;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMHY 10  
 ||| |||  
 Db 41 VVPSGIHY 48

## RESULT 2

F89892 carbamoyl-phosphate synthase large chain [imported] - Staphylococcus aureus (strain N315)  
 C:Species: Staphylococcus aureus  
 C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 22-Oct-2001

C:Accession: F89892  
 R:Kuroda, M.; Ohca, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.; Lancet 357, 1225-1240, 2001

A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.  
 A:Reference number: A89758; MUID:21311952; PMID:11418146

A:Accession: F89892

A>Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-1057 <KUR>

A:Cross-references: GB:BA000018; PID:g13701002; PIDN:BAB42298.1; GSPDB:GN00149

A:Experimental source: strain N315

C:Genetics:

A:Gene: PYRAB

C:Superfamily: carbamoyl-phosphate synthase (glutamine-hydrolyzing) large chain; biotin

Query Match 66.1%; Score 37; DB 2; Length 1057;  
 Best Local Similarity 60.0%; Pred. No. 39;  
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 EVVPXGMHYS 11  
 ||| |||

Db 190 EIVSNGLHYS 199

## RESULT 3

A42452

V1 protein - tobacco yellow dwarf virus (strain Australia)

C:Species: tobacco yellow dwarf virus

C:Date: 15-Jan-1993 #sequence\_revision 15-Jan-1993 #text\_change 08-Oct-1999

C:Accession: A42452

R:Morris, B.A.M.; Richardson, K.A.; Haley, A.; Zhan, X.; Thomas, J.E.

Virolgy 187, 633-642, 1992

A:Title: The nucleotide sequence of the infectious cloned DNA component of tobacco yellow

A:Reference number: A42452; MUID:92188538; PMID:1546458

A:Accession: A42452

A:Molecule type: DNA

A:Residues: 1-102 <MOR>

A:Cross-references: GB:MB1103; NID:g335283; PIDN:AAA47947.1; PID:g335284

Query Match 64.3%; Score 36; DB 2; Length 102;  
 Best Local Similarity 60.0%; Pred. No. 5;  
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 EVVPXGMHYS 11  
 ||| |||

Db 7 QVPSGINYS 16

## RESULT 4

AE2001

hypothetical protein alr1563 [imported] - Nostoc sp. (strain PCC 7120)

C:Species: Nostoc sp.

A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120

C:Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 30-Jun-2002

C:Accession: AE2001

R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.

DNA Res. 8, 205-213, 2001

A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium

A:Reference number: AB1807; MUID:21595285; PMID:11759840

A:Accession: AE2001

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-252 <KUR>

A:Cross-references: GB:BA000019; PIDN:BAB77929.1; PID:g17135383; GSPDB:GN00179

A:Experimental source: strain PCC 7120

C:Genetics:

A:Gene: alr1563

Query Match 64.3%; Score 36; DB 2; Length 252;  
 Best Local Similarity 50.0%; Pred. No. 13;  
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMHY 10  
 ||| |||

Db 235 EMIVPAGLHF 244

## RESULT 5

S69046

hypothetical protein YPL139c - yeast (Saccharomyces cerevisiae)

C:Species: Saccharomyces cerevisiae

C:Date: 22-Aug-1996 #sequence\_revision 06-Sep-1996 #text\_change 17-Mar-2000

C:Accession: S69046

R:Hall, J.; DePaulo, T.; Ahmed, A.; Bussey, H.; Fortin, N.; Friesen, J.D.; Storms, J.

submitted to the EMBL Data Library, December 1995

A:Description: The sequence of Saccharomyces cerevisiae chromosome XVI left arm.

A:Reference number: S69040

A:Accession: S69046

A:Molecule type: DNA

A:Residues: 1-460 <HAL>

A:Cross-references: EMBL:U43703; NID:g1244769; PIDN:AAB68221.1; PID:g1244776; MIPS:

C:Genetics:

A:Gene: SGD:UMEL

A:Cross-references: SGD:S0006060; MIPS:YPL139c

A:Map position: 16L

C:Superfamily: Saccharomyces cerevisiae transcription modulator WTM1

Query Match 64.3%; Score 36; DB 2; Length 460;  
 Best Local Similarity 62.5%; Pred. No. 25;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMHY 10  
 ||| |||

Db 85 IVPLGLHY 92

## RESULT 6

S38143

hypothetical protein YBL011w homolog YKR067w - yeast (Saccharomyces cerevisiae)

C:Species: Saccharomyces cerevisiae

C:Date: 03-May-1994 #sequence\_revision 03-May-1994 #text\_change 19-Apr-2002

C:Accession: S38143

R:van Vliet-Reedijk, J.C.; Planta, R.J.

submitted to the Protein Sequence Database, March 1994

A:Reference number: S38130

A:Accession: S38143

A:Molecule type: DNA

A:Residues: 1-743 <VAN>

A:Cross-references: EMBL:Z28292; NID:g486536; PIDN:CAA82146.1; PID:g486537; MIPS:YKR:

A:Experimental source: strain S288C

C:Genetics:

A:Gene: SGD:GPT2

A:Cross-references: SGD:S0001775

A:Map position: 11R

C:Keywords: transmembrane protein

Query Match 64.3%; Score 36; DB 2; Length 743;  
 Best Local Similarity 75.0%; Pred. No. 43;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMHY 10  
 ||| |||  
 Db 294 VPCGLHY 301  
 ||| |||

RESULT 7  
 DB2618  
 conserved hypothetical protein XF1950 [imported] - Xylella fastidiosa (strain 9a5c)  
 C:Species: Xylella fastidiosa  
 C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 20-Aug-2000  
 C:Accession: DB2618  
 R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequencing  
 Nature 406, 151-157, 2000  
 A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.  
 A:Reference number: AB2515; MUID:20365717; PMID:10910347  
 A:Note: for a complete list of authors see reference number A59328 below  
 A:Accession: DB2618  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-156 <SIM>  
 A:Cross-references: GB:AE004014; GB:AE003849; NID:g9107044; PIDN:AAF84752.1; GSPDB:GN001  
 A:Experimental source: strain 9a5c  
 R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A  
 Briones, M.R.S.; Bueno, M.R.P.; Camargo, L.E.A.; Carraro, D.M.; Carreir, H  
 as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.  
 submitted to GenBank, June 2000  
 A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm  
 J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig  
 chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, H  
 A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.  
 , F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A  
 Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak  
 A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir  
 M.; Tshakho, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z  
 A:Reference number: A59328  
 A:Contents: annotation  
 C:Genetics:  
 A:Gene: XF1950

Query Match 62.5%; Score 35; DB 2; Length 156;  
 Best Local Similarity 55.6%; Pred. No. 13;  
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXGMH 9  
 ||:| |||  
 Db 119 EEILPQGVH 127

RESULT 8  
 T02590  
 DNA binding protein BREBP-2 - common tobacco  
 C:Species: Nicotiana tabacum (common tobacco)  
 C:Date: 05-Mar-1999 #sequence\_revision 05-Mar-1999 #text\_change 21-Jul-2000  
 C:Accession: T02590  
 R:Ohme-Takagi, M.; Shinshi, H.  
 Plant Cell 7, 173-182, 1995  
 A:Title: Ethylene-inducible DNA binding proteins that interact with an ethylene responsi  
 A:Reference number: Z14671; MUID:95276459; PMID:7756828  
 A:Accession: T02590  
 A:Status: preliminary; translated from GB/EMBL/DBDJ  
 A:Molecule type: mRNA  
 A:Residues: 1-233 <OHM>  
 A:Cross-references: EMBL:D38126; NID:g790362; PIDN:BAA07324.1; PID:g1208498  
 A:Experimental source: strain BY4; tissue-type leaf

Query Match 62.5%; Score 35; DB 2; Length 233;  
 Best Local Similarity 60.0%; Pred. No. 19;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMHY 10  
 : ||| | ||  
 Db 90 QAVVPKGRHY 99

RESULT 9  
 H69194  
 GMP synthetase, subunit B - Methanobacterium thermoautotrophicum (strain Delta H)  
 C:Species: Methanobacterium thermoautotrophicum  
 C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 22-Oct-1999  
 C:Accession: H69194  
 R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge,  
 ; Qiu, D.; Spadafora, R.; Vicair, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwan  
 ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.  
 J. Bacteriol. 179, 7135-7155, 1997  
 A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: i  
 A:Reference number: A69000; MUID:98037514; PMID:9371463  
 A:Accession: H69194  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-311 <MTH>  
 A:Cross-references: GB:AE000850; GB:AE000666; NID:g2621794; PIDN:AAB85215.1; PID:g2  
 A:Experimental source: strain Delta H  
 C:Genetics:  
 A:Gene: MTH710  
 A:Start codon: GTG

Query Match 62.5%; Score 35; DB 2; Length 311;  
 Best Local Similarity 63.6%; Pred. No. 27;  
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMHY 11  
 |||| | |||  
 Db 219 EEVVGSLHES 229

RESULT 10  
 T24111  
 hypothetical protein R10D12.10 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
 C:Accession: T24111  
 R:Percy, C.  
 submitted to the EMBL Data Library, October 1996  
 A:Reference number: Z19842  
 A:Accession: T24111  
 A:Status: preliminary; translated from GB/EMBL/DBDJ  
 A:Molecule type: DNA  
 A:Residues: 1-425 <WIL>  
 A:Cross-references: EMBL:Z81109; PIDN:CAB03241.1; GSPDB:GN00023; CESP:R10D12.10  
 A:Experimental source: clone R10D12  
 C:Genetics:  
 A:Gene: CESP:R10D12.10  
 A:Map position: 5  
 A:Introns: 23/3; 56/3; 113/3; 257/2

Query Match 62.5%; Score 35; DB 2; Length 425;  
 Best Local Similarity 50.0%; Pred. No. 37;  
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMHY 10  
 |::| | | |  
 Db 335 EQIVPGGLQY 344

RESULT 11  
 G86430  
 TS18.1 protein - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 17-May-2002  
 C:Accession: G86430  
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Dewa  
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewa  
 ansen, N.F.; Hughes, B.; Huizar, L.  
 Nature 408, 816-820, 2000  
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; K

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziali, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
 A:Reference number: A86141; MUID:21016719; PMID:11130712  
 A:Accession: G86430  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-510 <STO>  
 A:Cross-references: GB:AE005172; NID:g4587512; PIDN:AAD25743.1; GSPDB:GN00141  
 C:Genetics:  
 A:Map position: 1  
 C:Superfamily: hexose phosphate transport protein uhpT

Query Match 62.5%; Score 35; DB 2; Length 510;  
 Best Local Similarity 60.0%; Pred. No. 45;  
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMHY 10  
 ||| | | | |  
 Db 12 EEVKPPGIHF 21

## RESULT 12

G69117  
 dipthine synthase - Methanobacterium thermoautotrophicum (strain Delta H)  
 C:Species: Methanobacterium thermoautotrophicum  
 C:Date: 05-Dec-1997 #sequence\_revision 03-Dec-1997 #text\_change 08-Oct-1999  
 C:Accession: G69117  
 R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.; Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N. ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.  
 J. Bacteriol. 179, 7135-7155, 1997  
 A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funct  
 A:Reference number: A69000; MUID:98037514; PMID:9371463  
 A:Accession: G69117  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-264 <MTH>  
 A:Cross-references: GB:AE000940; GB:AE000666; NID:g2623011; PIDN:AAB86340.1; PID:g262301  
 A:Experimental source: strain Delta H  
 C:Genetics:  
 A:Gene: MTH1874  
 C:Superfamily: Methanococcus jannaschii dipthine synthase

Query Match 60.7%; Score 34; DB 2; Length 264;  
 Best Local Similarity 62.5%; Pred. No. 35;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMHY 10  
 ||| | | | |  
 Db 235 VVPAGLHF 242

## RESULT 13

C75538  
 hypothetical protein - Deinococcus radiodurans (strain R1)  
 C:Species: Deinococcus radiodurans  
 C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 28-Jul-2000  
 C:Accession: C75538  
 R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma  
 S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.  
 Science 286, 1571-1577, 1999

A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.  
 A:Reference number: A75250; MUID:20036896; PMID:10567266  
 A:Accession: C75538  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-279 <WHI>  
 A:Cross-references: GB:AE001889; GB:AE000513; NID:g6457944; PIDN:AAF09867.1; PID:g645795  
 A:Experimental source: strain R1

C:Genetics:  
 A:Gene: DR0271  
 A:Map position: 1  
 C:Superfamily: Deinococcus radiodurans hypothetical protein DR0271

Query Match 60.7%; Score 34; DB 2; Length 279;  
 Best Local Similarity 75.0%; Pred. No. 38;  
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 VPXGMHYS 11  
 || | | | |  
 Db 100 VPLGRHYS 107

## RESULT 14

B75478  
 3-dehydroquininate synthase - Deinococcus radiodurans (strain R1)  
 C:Species: Deinococcus radiodurans  
 C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 17-Mar-2000  
 C:Accession: B75478  
 R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R. M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C. S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.  
 Science 286, 1571-1577, 1999  
 A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1  
 A:Reference number: A75250; MUID:20036896; PMID:10567266  
 A:Accession: B75478  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-350 <WHI>  
 A:Cross-references: GB:AE001932; GB:AE000513; NID:g6458481; PIDN:AAF10353.1; PID:g6  
 A:Experimental source: strain R1  
 C:Genetics:  
 A:Gene: DR0777  
 A:Map position: 1  
 C:Superfamily: 3-dehydroquininate synthase; 3-dehydroquininate synthase homology

Query Match 60.7%; Score 34; DB 2; Length 350;  
 Best Local Similarity 60.0%; Pred. No. 48;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 EVVPXGMHYS 11  
 || | | | | |  
 Db 252 EAVYGMHYA 261

## RESULT 15

T35025  
 probable DNA ligase - Streptomyces coelicolor  
 C:Species: Streptomyces coelicolor  
 C:Date: 05-Nov-1999 #sequence\_revision 05-Nov-1999 #text\_change 05-Nov-1999  
 C:Accession: T35025  
 R:Seeger, S.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A. submitted to the EMBL Data Library, June 1999  
 A:Reference number: Z21565  
 A:Accession: T35025  
 A:Status: preliminary; translated from GB/EMBL/DDBJ  
 A:Molecule type: DNA  
 A:Residues: 1-355 <SEE>  
 A:Cross-references: EMBL:AL079355; PIDN:CAB45581.1; GSPDB:GN00070; SCOEDB:SC4C6.17C  
 A:Experimental source: strain A3(2)  
 C:Genetics:  
 A:Gene: SCOEDB:SC4C6.17C

Query Match 60.7%; Score 34; DB 2; Length 355;  
 Best Local Similarity 71.4%; Pred. No. 49;  
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 VPXGMHY 10  
 : | | | | |  
 Db 20 IPPGMHY 26

## RESULT 16

E69086  
cell division protein - Methanobacterium thermoautotrophicum (strain Delta H)  
C:Species: Methanobacterium thermoautotrophicum  
C>Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 29-Sep-1999  
R:Accession: E69086  
R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.; Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwan, N.; S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.  
J. Bacteriol. 179, 7135-7155, 1997  
A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: function  
A:Reference number: A69000; MUID:98037514; PMID:9371463  
A:Accession: E69086  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-360 <MTH>  
A:Cross-references: GB:AE000923; GB:AE000666; NID:g2622766; PIDN:AAB86115.1; PID:g262277  
A:Experimental source: strain Delta H  
C:Genetics:  
A:Gene: MTH1642  
C:Superfamily: cell division protein MJ0174

Query Match 60.7%; Score 34; DB 2; Length 360;

Best Local Similarity 45.5%; Pred. No. 50;

Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMHYS 11

||| |||

Db 98 EDLVPMSGSHHT 108

## RESULT 17

C83903  
hypothetical protein BH2027 [imported] - Bacillus halodurans (strain C-125)  
C:Species: Bacillus halodurans  
C>Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 15-Jun-2001  
R:Accession: C83903  
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masul, N.; Fujii, F.; Hira  
Nucleic Acids Res. 28, 4317-4331, 2000  
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and  
A:Reference number: A83650; MUID:20512582; PMID:11058132  
A:Accession: C83903  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-425 <STO>  
A:Cross-references: GB:AP001514; GB:BA000004; NID:g10174613; PIDN:BA05746.1; GSPDB:GN00  
A:Experimental source: strain C-125  
C:Genetics:  
A:Gene: BH2027

Query Match 60.7%; Score 34; DB 2; Length 425;

Best Local Similarity 60.0%; Pred. No. 59;

Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMHY 10

||| |||

Db 403 EELLIEGMHY 412

## RESULT 18

S58132  
Slsl protein precursor - yeast (Yarrowia lipolytica)  
C:Species: Yarrowia lipolytica, Candida lipolytica  
C>Date: 13-Jan-1996 #sequence\_revision 01-Mar-1996 #text\_change 09-Sep-1997  
R:Accession: S58132  
R:Boisrame, A.; Beckerich, J.; Gaillardin, C.  
submitted to the EMBL Data Library, July 1995  
A:Description: Slslp, an endoplasmic reticulum component, is involved in the protein tra  
A:Reference number: S58132  
A:Accession: S58132  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-426 <BOI>

A:Cross-references: EMBL:Z50154; NID:g1052827; PID:g1052828

Query Match 60.7%; Score 34; DB 2; Length 426;

Best Local Similarity 44.4%; Pred. No. 59;

Matches 4; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXGMH 9

||| |||

Db 52 DQVAPAGLH 60

## RESULT 19

T28717  
hypothetical protein F10D2.3 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 29-Oct-1999  
R:Accession: T28717  
R:Graves, T.; Wohlmann, P.; Gillam, B.  
submitted to the EMBL Data Library, September 1997  
A:Description: The sequence of C. elegans cosmid F10D2.  
A:Reference number: 220515  
A:Accession: T28717  
A:Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-495 <GRA>  
A:Cross-references: EMBL:AF022972; PIDN:AAC48234.1; GSPDB:GN00023; CESP:F10D2.3  
A:Experimental source: strain Bristol N2; clone F10D2  
C:Genetics:  
A:Gene: CESP:F10D2.3  
A:Map position: 5  
A:Introns: 37/2; 90/1; 113/1; 183/3; 356/3; 380/3; 428/2

Query Match 60.7%; Score 34; DB 2; Length 495;

Best Local Similarity 50.0%; Pred. No. 70;

Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMHY 10

||| |||

Db 218 ENIVPTGKHH 227

## RESULT 20

AF3286  
ATP-dependent DNA helicase BMEI0275 [imported] - Brucella melitensis (strain 16M)  
C:Species: Brucella melitensis  
C>Date: 01-Feb-2002 #sequence\_revision 01-Feb-2002 #text\_change 01-Feb-2002  
R:Accession: AF3286  
R:Delvecchio, V.G.; Kapatal, V.; Redkar, R.J.; Patra, G.; Mujar, C.; Los, T.; Iva  
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002  
A:Title: The genome sequence of the facultative intracellular pathogen Brucella me  
A:Reference number: AD3252; PMID:11756688  
A:Accession: AF3286  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1028 <KUR>  
A:Cross-references: GB:AE008917; PIDN:AAL51457.1; PID:g17982167; GSPDB:GN00190  
A:Experimental source: strain 16M  
C:Genetics:  
A:Gene: BMEI0275  
A:Map position: 1

Query Match 60.7%; Score 34; DB 2; Length 1028;

Best Local Similarity 54.5%; Pred. No. 1.5e+02;

Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMHYS 11

||| |||

Db 76 EKIVPPGARYS 86

## RESULT 21

S54619

```

Query Match      58.9%;   Score 33;   DB 2;   Length 441;
Best Local Similarity 54.5%;   Pred. No. 98;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY      1  EEVVPXGMHYS 11
          | | | | : | | :
Db      417  ETVPPTFIHYN 427

```

RESULT 24  
G71542

probable amino acid transporter - Chlamydia trachomatis (serotype D, strain UW3/Cx)  
C:Species: Chlamydia trachomatis  
C:Date: 13-Sep-1998 #sequence\_revision 13-Sep-1998 #text\_change 19-May-2000  
C:C.Accession: G71542  
R:Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchell, A.P.  
S:Science 282, 754-759, 1998  
A:Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia  
A:Reference number: A71570; MUID:99000809; PMID:9784136  
A:C.accession: G71542  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-466 <ARN>  
A:Cross-references: GB:AE001295; GB:AE001273; NID:g3328617; PIDN:AAC67808.1; PID:g3328617  
A:Experimental source: serotype D, strain UW-3/Cx  
C:Genetics:  
A:Gene: xsaA  
C:Superfamily: L-lysine transport protein

amino acid antiporter TC0488 [imported] - Chlamydia muridarum (Strain Nigg)  
 ;Species: Chlamydia muridarum, Chlamydia trachomatis MoPn  
 ;Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 02-Jun-2000  
 ;Accession: H81697  
 ;Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hick  
 ;C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzbu  
 ;Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR3  
 ;Reference number: A81500; MUID:20150255; PMID:10684935  
 ;Accession: H81697  
 ;Status: preliminary  
 ;Molecule type: DNA  
 ;Residues: 1-466 <TEF>  
 ;Cross-references: GB:AE002317; GB:AE002160; NID:g7190522; PIDN:AAF39334.1; PID:g71  
 ;Experimental source: strain Nigg (MoPn)  
 ;Genetics:  
 ;Gene: TC0488  
 ;Superfamily: L-lysine transport protein

Query Match 58.9%; Score 33; DB 2; Length 466;  
Best Local Similarity 62.5%; Fred. No. 1e+02;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 4 VPXGMHYS 11  
: | | | | |  
Db 453 IPFGMYYS 460

Search completed: June 10, 2003, 13:49:11  
Job time : 13.2143 secs





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2003, 13:25:04 ; Search time 4.5 Seconds  
(without alignments)  
101.387 Million cell updates/sec

Title: US-09-909-164-7  
Perfect score: 56  
Sequence: 1 BEVVPXGMHYS 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues  
Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 75 summaries

Database: SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	38	67.9	1058	1	CARB_FUSNN
2	37	66.1	1057	1	CARB_STAAN
3	37	66.1	1057	1	CARB_STAAN
4	36	64.3	102	1	Y1LK_TYDVA
5	36	64.3	460	1	UMEL_YEAST
6	36	64.3	743	1	YK47_YEAST
7	35	62.5	227	1	ID11_MESAU
8	35	62.5	308	1	GAAB_METH
9	34	60.7	426	1	SLSL_YARLI
10	33	58.9	513	1	PHSL_DESBA
11	33	58.9	627	1	MUTL_BACSU
12	33	58.9	1188	1	KPBA_CAEEL
13	33	58.9	1396	1	V1TF_BPT5
14	32	57.1	126	1	TKNK_BOVIN
15	32	57.1	267	1	RR2_CHLVU
16	32	57.1	288	1	CGD2_RAT
17	32	57.1	289	1	CGD2_HUMAN
18	32	57.1	289	1	CGD2_MOUSE
19	32	57.1	291	1	CGD1_BRARE
20	32	57.1	291	1	CGD1_XENLA
21	32	57.1	291	1	CGD2_CHICK
22	32	57.1	291	1	CGD2_XENLA
23	32	57.1	292	1	CGD1_CHICK
24	32	57.1	292	1	CGD3_HUMAN
25	32	57.1	295	1	CGD1_HUMAN
26	32	57.1	295	1	CGD1_MOUSE
27	32	57.1	295	1	CGD1_RAT
28	32	57.1	341	1	HYPE_AZOVI
29	32	57.1	353	1	T2BA_BACAR
30	32	57.1	759	1	SULH_YEAST
31	32	57.1	877	1	SULH_SCHPO
32	32	57.1	1401	1	RPOC_VIBCH
33	32	57.1	2717	1	ZEP1_HUMAN

34	31.5	56.2	847	1	CD22_HUMAN
35	31	55.4	124	1	REV_SIVCZ
36	31	55.4	130	1	SZ05_RAT
37	31	55.4	276	1	Y939_METJA
38	31	55.4	319	1	YHAL_CRIPI
39	31	55.4	331	1	RL3_ARCFU
40	31	55.4	363	1	ALFB_HAEIN
41	31	55.4	450	1	Y325_HAEIN
42	31	55.4	529	1	ENP3_HUMAN
43	31	55.4	609	1	PTBA_BACSU
44	31	55.4	674	1	DCMB_MOOTH
45	31	55.4	735	1	ETRI_BRAOL
46	31	55.4	738	1	ETRI_ARATH
47	31	55.4	788	1	CY14_NEUCR
48	31	55.4	906	1	CENC_MOUSE
49	31	55.4	1374	1	YQS6_CAEEL
50	31	55.4	1378	1	RON_MOUSE
51	31	55.4	1394	1	LTBS_HUMAN
52	31	55.4	1429	1	EXPA_DROME
53	31	55.4	1498	1	Y1A9_CLOAB
54	31	55.4	1595	1	LTBL_HUMAN
55	31	55.4	1712	1	LTBL_RAT
56	30.5	54.5	73	1	IF1_CHLPN
57	30.5	54.5	73	1	IF1_CHLTR
58	30.5	54.5	492	1	CATL_ARATH
59	30	53.6	17	1	ND4M_TRIRU
60	30	53.6	121	1	TKNK_HUMAN
61	30	53.6	152	1	Y16D_BPT4
62	30	53.6	172	1	YKKB_BACSU
63	30	53.6	185	1	PAGC_SALTY
64	30	53.6	212	1	MSRA_VIBCH
65	30	53.6	219	1	SSL_DROME
66	30	53.6	223	1	COAT_CTV36
67	30	53.6	232	1	SCOA_HELPJ
68	30	53.6	232	1	SCOA_HELPY
69	30	53.6	298	1	YA52_HAEIN
70	30	53.6	331	1	LDHC_HUMAN
71	30	53.6	358	1	YVAA_BACSU
72	30	53.6	363	1	PIT1_CHICK
73	30	53.6	370	1	PIT1_MELGA
74	30	53.6	396	1	DHH_HUMAN
75	30	53.6	396	1	DHH_MOUSE

## ALIGNMENTS

## RESULT 1

CARB_FUSNN	STANDARD;	PRT; 1058 AA.
ID	CARB_FUSNN	
AC	Q8RG86;	
DT	15-JUN-2002 (Rel. 41, Created)	
DT	15-JUN-2002 (Rel. 41, Last sequence update)	
DT	15-JUN-2002 (Rel. 41, Last annotation update)	
DE	Carbamoyl-phosphate synthase large chain (EC 6.3.5.5) (Carbamoyl-phosphate synthase ammonia chain).	
GN	CARB OR FN0422.	
OS	Fusobacterium nucleatum (subsp. nucleatum).	
OC	Bacteria; Fusobacteria; Fusobacterium.	
OX	NCBI_TaxID=76856;	
RP	[1]	
RC	SEQUENCE FROM N.A.	
RC	STRAIN=21886394; PubMed=11889109;	
RA	Kapatal V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A.,	
RA	Bhattacharyya A., Bartman A., Gardner W., Grechkin G., Zhu L.,	
RA	Vasileva O., Chu L., Kogan Y., Chaga O., Goltzman E., Bernal A.,	
RA	Larsen N., D'Souza M., Walunas T., Pusch G., Haselkorn R.,	
RA	Fontstein N., Kyrpides N., Overbeek R.;	
RT	"Genome sequence and analysis of the oral bacterium Fusobacterium	
RT	nucleatum strain ATCC 25586."	
RL	J. Bacteriol. 184:2005-2018(2002).	
CC	-1- CATALYTIC ACTIVITY: 2 ATP + L-glutamine + CO(2) + H(2)O -> 2 ADP +	

P2073	homo sapien
P17280	chimpanzee
P97885	rattus norv
Q58349	methanococ
P10941	crithonec
O28354	archaeoglob
P22210	ovis aries
P46640	haemophilus
O75355	homo sapien
P40739	bacillus su
P27989	moorella th
O49230	brassica ol
P49333	arabidopsis
P23622	neurospora
P49452	mus musculu
Q09541	caenorhabdi
Q62190	mus musculu
P22064	homo sapien
Q07436	drosophila
Q04351	clostridium
Q14766	homo sapien
Q00918	rattus norv
Q92948	chlamydia p
O84325	chlamydia t
Q96528	arabidopsis
Q36834	trichophyto
Q9uH0	homo sapien
P22917	bacterioph
P49855	bacillus su
P23988	salmonella
Q3Kp30	vibrio chol
Q24536	drosophila
Q00686	citrus tris
Q92le3	helicobacte
P56006	helicobacte
P45008	haemophilus
P07864	homo sapien
Q32223	bacillus su
Q9Ygl7	gallus gall
Q05749	meleagris g
O43323	homo sapien
O61488	mus musculu

phosphate + L-glutamate + carbamoyl phosphate.  
 -1- COPACTOR: Binds three manganese ions (By similarity).  
 -1- PATHWAY: Arginine biosynthesis.  
 -1- SUBUNIT: Composed of two chains; first step.  
 promotes the hydrolysis of glutamine to ammonia, which is used by the large (or ammonia) chain to synthesize carbamoyl phosphate (By similarity).  
 -1- SIMILARITY: BELONGS TO THE CARB FAMILY.  
 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 EMBL; AB010554; AAL94625.1; ALT\_INIT.  
 DR InterPro: IPR005483; CPase\_L.  
 DR InterPro: IPR005479; CPase\_L\_D2.  
 DR InterPro: IPR005480; CPase\_L\_D3.  
 DR InterPro: IPR005481; CPase\_L\_N.  
 DR InterPro: IPR004362; MGS-like.  
 DR Pfam: PF00289; CPase\_L\_chain; 2.  
 DR Pfam: PF02786; CPase\_L\_D2; 2.  
 DR Pfam: PF02787; CPase\_L\_D3; 1.  
 DR Pfam: PF02142; MGS; 1.  
 DR PRINTS; PR00098; CPASE.  
 DR PROSITE; PS00866; CPASE\_1; 2.  
 DR PROSITE; PS00867; CPASE\_2; 2.  
 KW Arginine biosynthesis; Pyrimidine biosynthesis; Ligase; Repeat;  
 ATP-binding; Manganese; Complete proteome.  
 FT DOMAIN 1 401 CARBOXYPHOSPHATE SYNTHETIC DOMAIN.  
 FT DOMAIN 402 546 OLIGOMERIZATION DOMAIN.  
 FT DOMAIN 547 929 CARBAMOYL PHOSPHATE SYNTHETIC DOMAIN.  
 FT REPEAT 1 546 ALLOSTERIC DOMAIN.  
 FT NP\_BIND 153 210 ATP (POTENTIAL).  
 FT NP\_BIND 302 352 ATP (POTENTIAL).  
 FT METAL 284 284 MANGANESE 1 (BY SIMILARITY).  
 FT METAL 298 298 MANGANESE 1 AND 2 (BY SIMILARITY).  
 FT METAL 300 300 MANGANESE 2 (BY SIMILARITY).  
 FT METAL 820 820 MANGANESE 3 (BY SIMILARITY).  
 FT METAL 832 832 MANGANESE 3 (BY SIMILARITY).  
 SQ SEQUENCE 1058 AA; ED7037AF77C1E39F CRC64;  
 Query Match 67.9%; Score 38; DB 1; Length 1058;  
 Best Local Similarity 60.0%; Pred. No. 10;  
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
 QY 2 EVVPGMHYS 11  
 Db 190 EIVPGLNYS 199  
 RESULT 2  
 CARB\_STAAM STANDARD; PRT; 1057 AA.  
 AC Q99UR5;  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Carbamoyl-phosphate synthase large chain (EC 6.3.5.5) (Carbamoyl-phosphate synthetase ammonia chain).  
 OS CARB OR PYRAB OR SAV1203 OR SA1046.  
 GN Staphylococcus aureus (strain N315).  
 OS Staphylococcus aureus (strain N315).  
 OC Bacteria; Firmicutes; Bacillales; Staphylococcus.  
 NCBI\_TaxID=158878, 158879;  
 RN [1]  
 RP SEQUENCE FROM N.A.

STRAIN=MU50 / ATCC 700699, and N315;  
 MEDLINE=21311952; PubMed=11418146;  
 RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I., Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T., Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A., Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C., Sekimizu K., Hirakawa H., Kihara S., Goto S., Yabuzaki J., Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T., Hattori M., Ogasawara N., Hayashi H., Hiratsuku K.;  
 RT "Whole genome sequencing of methicillin-resistant Staphylococcus aureus.";  
 RL Lancet 357:1225-1240(2001).  
 CC -1- CATALYTIC ACTIVITY: 2 ATP + L-glutamine + CO(2) + H(2)O = 2 ADP + phosphate + L-glutamate + carbamoyl phosphate.  
 CC -1- COPACTOR: Binds three manganese ions (By similarity).  
 CC -1- PATHWAY: Arginine biosynthesis.  
 CC -1- SUBUNIT: Pyrimidine biosynthesis; first step.  
 promotes the hydrolysis of glutamine to ammonia, which is used by the large (or ammonia) chain to synthesize carbamoyl phosphate (By similarity).  
 -1- SIMILARITY: BELONGS TO THE CARB FAMILY.  
 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 EMBL; AP003361; BAB57365.1;  
 EMBL; AP003132; BAB42298.1;  
 HSPP; P00968; 1CS0.  
 DR InterPro: IPR005483; CPase\_L.  
 DR InterPro: IPR005479; CPase\_L\_D2.  
 DR InterPro: IPR005480; CPase\_L\_D3.  
 DR InterPro: IPR005481; CPase\_L\_N.  
 DR InterPro: IPR004362; MGS-like.  
 DR Pfam: PF00289; CPase\_L\_chain; 2.  
 DR Pfam: PF02786; CPase\_L\_D2; 2.  
 DR Pfam: PF02787; CPase\_L\_D3; 1.  
 DR Pfam: PF02142; MGS; 1.  
 DR PRINTS; PR00098; CPASE.  
 DR PROSITE; PS00866; CPASE\_1; 2.  
 DR PROSITE; PS00867; CPASE\_2; 2.  
 KW Arginine biosynthesis; Pyrimidine biosynthesis; Ligase; Repeat;  
 ATP-binding; Manganese; Complete proteome.  
 FT DOMAIN 1 401 CARBOXYPHOSPHATE SYNTHETIC DOMAIN.  
 FT DOMAIN 402 546 OLIGOMERIZATION DOMAIN.  
 FT DOMAIN 547 929 CARBAMOYL PHOSPHATE SYNTHETIC DOMAIN.  
 FT REPEAT 1 546 ALLOSTERIC DOMAIN.  
 FT NP\_BIND 153 1057 ATP (POTENTIAL).  
 FT NP\_BIND 153 210 ATP (POTENTIAL).  
 FT METAL 302 352 MANGANESE 1 (BY SIMILARITY).  
 FT METAL 284 284 MANGANESE 1 AND 2 (BY SIMILARITY).  
 FT METAL 298 298 MANGANESE 2 (BY SIMILARITY).  
 FT METAL 300 300 MANGANESE 3 (BY SIMILARITY).  
 FT METAL 820 820 MANGANESE 3 (BY SIMILARITY).  
 FT METAL 832 832 MANGANESE 3 (BY SIMILARITY).  
 SQ SEQUENCE 1057 AA; 117171 MW; E3E179EF0591F0F8 CRC64;  
 Query Match 66.1%; Score 37; DB 1; Length 1057;  
 Best Local Similarity 60.0%; Pred. No. 16;  
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
 QY 2 EVVPGMHYS 11  
 Db 190 EIVSGLNYS 199  
 RESULT 3

CARB\_STAAW STANDARD; PRT; 1057 AA.

AC P58940;

DT 15-JUN-2002 (Rel. 41, Created)

DT 15-JUN-2002 (Rel. 41, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Carbamoyl-phosphate synthase large chain (EC 6.3.5.5) (Carbamoyl-phosphate synthetase ammonia chain).

GN CARB OR PYRAB OR MW1086.

OS Staphylococcus aureus (strain MW2).

OC Bacteria; Firmicutes; Bacillales; Staphylococcus.

OX NCBI\_TaxID=196620;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=22040717; PubMed=12044378;

RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A., Nagai Y., Iwama N., Asano K., Naimi T., Kuroda H., Cui L., Yamamoto K., Hiramatsu K.;

RA "Genome and virulence determinants of high virulence community-acquired MRSA";

RT Lancet 359:1819-1827(2002).

RL -|- CATALYTIC ACTIVITY: 2 ATP + L-glutamine + CO(2) + H(2)O = 2 ADP + phosphate + L-glutamate + carbamoyl phosphate.

CC -|- COFACTOR: Binds three manganese ions (By similarity).

CC -|- PATHWAY: Arginine biosynthesis; first step.

CC -|- SUBUNIT: Pyrimidine biosynthesis; the small (or glutamine) chain promotes the hydrolysis of glutamine to ammonia, which is used by the large (or ammonia) chain to synthesize carbamoyl phosphate (By similarity).

CC -|- SIMILARITY: BELONGS TO THE CARB FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

EMBL; AP004825; BAB94951.1; -.

DR PROSITE; PS00866; CPSASE\_1; 2.

DR PROSITE; PS00867; CPSASE\_2; 2.

KW Arginine biosynthesis; Pyrimidine biosynthesis; Ligase; Repeat;

KW ATP-binding; Manganese.

FT DOMAIN 1 401 CARBOXYPHOSPHATE SYNTHETIC DOMAIN.

FT DOMAIN 402 546 OLIGOMERIZATION DOMAIN.

FT DOMAIN 547 929 CARBAMOYL PHOSPHATE SYNTHETIC DOMAIN.

FT DOMAIN 930 1057 ALLOSTERIC DOMAIN.

FT REPEAT 1 546

FT REPEAT 547 1057

FT NP\_BIND 153 210 ATP (POTENTIAL).

FT NP\_BIND 302 352 ATP (POTENTIAL).

FT METAL 284 284 MANGANESE 1 (BY SIMILARITY).

FT METAL 298 298 MANGANESE 1 AND 2 (BY SIMILARITY).

FT METAL 300 300 MANGANESE 2 (BY SIMILARITY).

FT METAL 820 820 MANGANESE 3 (BY SIMILARITY).

FT METAL 832 832 MANGANESE 3 (BY SIMILARITY).

SQ SEQUENCE 1057 AA; 117185 MW; D8E3B09F9BC6F152 CRC64;

Query Match 66.1%; Score 37; DB 1; Length 1057;

Best Local Similarity 60.0%; Pred. No. 16;

Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 EVVPGXGMHYS 11

Db 190 EIVSNGLHYS 199

RESULT 4

Y1LK\_TYDVA STANDARD; PRT; 102 AA.

ID Y1LK\_TYDVA

AC P31619;

DT 01-JUL-1993 (Rel. 26, Created)

DT 01-JUL-1993 (Rel. 26, Last sequence update)

DT 01-OCT-1993 (Rel. 27, Last annotation update)

DE Hypothetical 11.2 kDa protein.

GN VI.

OS Tobacco yellow dwarf virus (strain Australia) (TYDV).

OC Viruses; ssDNA viruses; Geminiviridae; Mastrevirus.

OX NCBI\_TaxID=31599;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=92188538; PubMed=1546458;

RA Morris B.A.M., Richardson K.A., Haley A., Zhan X., Thomas J.E.;

RA "The nucleotide sequence of the infectious cloned DNA component of tobacco yellow dwarf virus reveals features of geminiviruses infecting monocotyledonous plants.";

RT Virology 187:633-642(1992).

RL -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

EMBL; M81103; AAA47947.1; -.

DR PIR; A42452; A42452.

DR InterPro; IPR002621; Gemini\_mov.

DR Pfam; PF01708; Gemini\_mov; 1.

KW Hypothetical protein.

SQ SEQUENCE 102 AA; 11178 MW; A40ECF1E0AF55B67 CRC64;

Query Match 64.3%; Score 36; DB 1; Length 102;

Best Local Similarity 60.0%; Pred. No. 2.3;

Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 EVVPGXGMHYS 11

Db 7 QVVPSSGINS 16

RESULT 5

UMEL\_YEAST STANDARD; PRT; 460 AA.

ID UMEL\_YEAST

AC Q03010; P87330;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Melosis negative regulator UMEL.

GN UMEL OR WTM3 OR YPL139C OR LPI7C.

OS Saccharomyces cerevisiae (Baker's yeast).

OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; Saccharomycetaceae; Saccharomyces.

OX NCBI\_TaxID=4932;

RN [1]

RP SEQUENCE FROM N.A.

RX STRAIN=A364A;

RA Malloy M.J., Strich R.;

RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RX STRAIN=S288C / AB972;

RX MEDLINE=97313271; PubMed=9169875;

RA Bussey H., Storms R.K., Ahmed A., Albermann K., Allen E., Ansoorge W., Araujo R., Aparicio A., Barrell B.G., Badcock K., Benes V., Botstein D., Bowman S., Bruckner M., Carpenter J., Cherry J.M., Chung E., Churcher C.M., Coster F., Davis K., Davis R.W., Dietrich F.S., Delius H., DiPaolo T., Dubois E., Duesterhoeft A., Duncan M., Floeth M., Fortin N., Friesen J.D., Fritz C., Goffeau A., Hall J., Hebling U., Heumann K., Hilbert H., Hallier L., Hunnicke-Smith S., Hyman R., Johnston M., Kalman S., Kleine K., Komp C., Kurdi O., Lashkari D., Lew H., Lin A., Lin D., Louis E.J., Marathe R., Messenguy F., Mewes H.-W., Mirtipati S., Moestl D.,

RA Mueller-Auer S., Namath A., Nentwich U., Oefner P., Pearson D.,  
 RA Petel F.X., Pohl T.M., Purnelle D., Schafer M., Scharfe M.,  
 RA Scherens B., Schramm S., Schroeder M., Sdicu A.M., Tettelin H.,  
 RA Urrestarazu L.A., Ushinsky S., Vierendeels F., Vissers S., Voss H.,  
 RA Walsh S.V., Wambutt R., Wang Y., Wedler E., Wedler H., Winnett E.,  
 RT Zhong W.W., Zollner A., Vo D.H., Hani J.,  
 RL The nucleotide sequence of Saccharomyces cerevisiae chromosome XVI.;  
 Nature 387:103-105 (1997).  
 CC -!- FUNCTION: TRANSCRIPTIONAL MODULATOR WITH ROLES IN MEIOTIC  
 CC REGULATION AND SILENCING. NEGATIVE REGULATOR OF MEIOTIC  
 CC -!- SIMILARITY: CONTAINS 4 WD REPEATS (TRP-ASP DOMAINS).  
 CC -!- SIMILARITY: STRONG, TO YEAST WTM1 AND WTM2.  
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; U10280; AAB40937.1; -;  
 DR EMBL; U43703; AAB68221.1; -;  
 DR TRANSFAC; T04309; -;  
 DR SGD; S0006060; UME1.  
 DR InterPro; IPR001680; WD40.  
 DR Pfam; PF00400; WD40; 3.  
 DR SMART; SM00320; WD40; 3.  
 DR PROSITE; PS00678; WD\_REPEATS\_1; FALSE NEG.  
 DR PROSITE; PS50082; WD\_REPEATS\_2; FALSE NEG.  
 DR PROSITE; PS50294; WD\_REPEATS\_REGION; FALSE NEG.  
 KW Transcription regulation; Meiosis; Repeat; WD repeat.  
 FT REPEAT 233 271 WD 1.  
 FT REPEAT 276 316 WD 2.  
 FT REPEAT 339 379 WD 3.  
 FT REPEAT 411 451 WD 4.  
 FT REPEAT 460 460 WD 4.  
 SQ SEQUENCE 460 AA; 51022 MW; AA6F60448B7BCBA9 CRC64;  
 Query Match 64.3%; Score 36; DB 1; Length 460;  
 Best Local Similarity 62.5%; Pred. No. 11;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 QY 3 VVPXGMHY 10  
 DB 85 IVPLGLHY 92  
 RESULT 6  
 YK47\_YEAST STANDARD; PRT; 743 AA.  
 AC P36148;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 01-NOV-1997 (Rel. 35, Last annotation update)  
 DE Hypothetical 83.6 kDa protein in CCp1-MET1 intergenic region.  
 GN YKR067W.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
 OX NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=S288C;  
 RA van Vliet-Reedijk J.C., Planta R.J.;  
 RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (probable).  
 CC -!- SIMILARITY: STRONG, TO YEAST YBL011W.  
 CC -----  
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; Z28292; CAA82146.1; -;  
 DR PIR; S38143; S38143.  
 DR SGD; S0001775; YKR067W.  
 DR InterPro; IPR002123; Acyltransferase.  
 KW Hypothetical protein; Transmembrane.  
 FT TRANSMEM 31 55 POTENTIAL.  
 FT TRANSMEM 69 85 POTENTIAL.  
 FT TRANSMEM 502 524 POTENTIAL.  
 FT TRANSMEM 539 555 POTENTIAL.  
 SQ SEQUENCE 743 AA; 83644 MW; 84B9946E56B82F15 CRC64;  
 Query Match 64.3%; Score 36; DB 1; Length 743;  
 Best Local Similarity 75.0%; Pred. No. 18;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 3 VVPXGMHY 10  
 DB 294 VVPCGLHY 301  
 RESULT 7  
 ID11\_MESAU STANDARD; PRT; 227 AA.  
 AC O35586;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Isopentenyl-diphosphate delta-isomerase 1 (EC 5.3.3.2) (IPP isomerase  
 DE 1) (Isopentenyl pyrophosphate isomerase 1) (IPP1).  
 GN ID11.  
 OS Mesocricetus auratus (Golden hamster).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;  
 OC Mesocricetus.  
 OX NCBI\_TaxID=10036;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97373600; PubMed=9228075;  
 RA Paton V.G., Shackelford J.E., Kilsans S.K.;  
 RT "Cloning and subcellular localization of hamster and rat isopentenyl  
 RT diphosphate dimethylallyl diphosphate isomerase. A PtsI motif targets  
 RT the enzyme to peroxisomes."  
 RL J. Biol. Chem. 272:18945-18950 (1997).  
 CC -!- FUNCTION: CATALYZES THE 1,3-ALLYLIC REARRANGEMENT OF THE  
 CC HOMALLYLIC SUBSTRATE ISOPENTENYL (IPP) TO ITS HIGHLY  
 CC ELECTROPHILIC ALLYLIC ISOMER, DIMETHYLLALLYL DIPHOSPHATE (DMAPP).  
 CC -!- CATALYTIC ACTIVITY: Isopentenyl diphosphate -> dimethylallyl  
 CC diphosphate.  
 CC -!- COFACTOR: REQUIRES MAGNESIUM FOR ACTIVITY.  
 CC -!- PATHWAY: ISOPRENOID BIOSYNTHETIC PATHWAY WHOSE END PRODUCTS  
 CC INCLUDE DOLICHOLS, VITAMINS A, D, E, AND K, STEROID HORMONES,  
 CC CAROTENOIDS BILE ACIDS AND CHOLESTEROL.  
 CC -!- SUBCELLULAR LOCATION: Peroxisomal.  
 CC -!- SIMILARITY: BELONGS TO THE IPP ISOMERASE TYPE 1 FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AF003836; AAC53283.1; -;  
 DR InterPro; IPR002667; IPP\_isomerase.  
 DR InterPro; IPR000086; NUDIX\_hydrolase.  
 DR Pfam; PF00293; NUDIX; 1.  
 DR ProDom; PD004109; IPP\_isomerase; 1.  
 KW Isomerase; Isoprene biosynthesis; Cholesterol biosynthesis;  
 KW Sterol biosynthesis; Peroxisome; Magnesium.

```

FT ACT_SITE 86 86 BY SIMILARITY.
FT ACT_SITE 148 148 BY SIMILARITY.
FT SITE 225 227 MICROBODY TARGETING SIGNAL.
SQ SEQUENCE 227 AA; F500A6586385E803 CRC64;

Query Match 62.5%; Score 35; DB 1; Length 227;
Best Local Similarity 70.0%; Pred. No. 8.5;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMHY 10
    ||| | |||
    121 EEVDPMHY 130

Db

RESULT 8
GAAB_METH STANDARD; PRT; 308 AA.
AC O26806;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE GMP synthase [glutamine-hydrolyzing] subunit B (EC 6.3.5.2) (GMP
DE synthetase).
GN GAAB OR MTH710.
OS Methanobacterium thermoautotrophicum.
OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
OC Methanobacteriaceae; Methanothermobacter.
OX NCBI_TaxID=187420;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Delta H;
RX MEDLINE=98037514; PubMed=9371463;
RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,
RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
RA Harrison D., Hoang L., Keagle P., Lumm W., Pothier B., Qiu D.,
RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
RA Jiواني N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,
RA McDougall S., Shimer G., Goyal A., Pietrowski S., Church G.M.,
RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;
RT "Complete genome sequence of Methanobacterium thermoautotrophicum
RT deltaH: functional analysis and comparative genomics."
RL J. Bacteriol. 179:7135-7155(1997).
CC -1- CATALYTIC ACTIVITY: ATP + xanthosine 5'-phosphate + L-glutamine +
CC H(2)O = AMP + diphosphate + GMP + L-glutamate.
CC -1- PATHWAY: GMP biosynthesis.
CC -1- SUBUNIT: HETERODIMER COMPOSED OF A GLUTAMINE AMIDOTRANSFERASE
CC SUBUNIT (A) AND A GMP SYNTHASE SUBUNIT (B) (POTENTIAL).
CC -1- SIMILARITY: BELONGS TO THE GMP SYNTHASE FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; AE000850; AAB85215.1; ALT_INIT.
CC HSPSP; P04079; 1GPM.
CC InterPro; IPR001674; GMP_synt_C.
CC Pfam; PF00958; GMP_synt_C; 1.
CC TIGRFAMs; TIGR00884; guaA.Cterm; 1.
CC Ligase; GMP biosynthesis; Purine biosynthesis; ATP-binding;
CC Complete proteome. 184 GMP-BINDING (BY SIMILARITY).
FT DOMAIN 33 184
FT NP_BIND 29 35 ATP (BY SIMILARITY).
SQ SEQUENCE 308 AA; 34403 MW; F2DCF6ED202CAEC1 CRC64;

Query Match 62.5%; Score 35; DB 1; Length 308;
Best Local Similarity 63.6%; Pred. No. 12;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMHY 11

```

```

Db 216 EEVVEGLHES 226
    |||| | | |
    216 EEVVEGLHES 226

RESULT 9
SLSL_YARLI STANDARD; PRT; 426 AA.
ID SLSL_YARLI
AC Q99158;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE SLS1 protein precursor.
GN SLS1.
OS Yarrowia lipolytica (Candida lipolytica).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Dipodascaceae; Yarrowia.
OX NCBI_TaxID=4952;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 20460 / W29;
RX MEDLINE=96216076; PubMed=8662639;
RA Boisrame A., Beckerich J.-M., Gaillardin C.;
RT "Sls1, an endoplasmic reticulum component, is involved in the
RT protein translocation process in the yeast Yarrowia lipolytica."
RL J. Biol. Chem. 271:11668-11675(1996).
CC -1- FUNCTION: INVOLVED IN THE PROTEIN TRANSLOCATION PROCESS. MAY
CC INTERACT DIRECTLY WITH TRANSLATING POLYPEPTIDES TO FACILITATE
CC THEIR TRANSFER AND/OR HELP THEIR FOLDING IN THE ER. IT IS NOT
CC REQUIRED FOR VIABILITY BUT IS ESSENTIAL FOR OPTIMAL GROWTH AT
CC ELEVATED TEMPERATURES.
CC -1- SUBCELLULAR LOCATION: Endoplasmic reticulum lumen.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; Z50154; CAA90516.1; -.
CC InterPro; IPR000886; ER_target.
CC PROSITE; PS00014; ER_TARGET; 1.
CC Endoplasmic reticulum; Signal.
FT SIGNAL 1 17 POTENTIAL.
FT CHAIN 18 426 SLS1 PROTEIN.
FT SITE 423 426 PREVENT SECRETION FROM ER (POTENTIAL).
SQ SEQUENCE 426 AA; 47201 MW; 0ACD7EF17540B8E2 CRC64;

Query Match 60.7%; Score 34; DB 1; Length 426;
Best Local Similarity 44.4%; Pred. No. 26;
Matches 4; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXGMH 9
    : | : | |
    52 DQVTPAGLH 60

Db

RESULT 10
PHSL_DESBA STANDARD; PRT; 513 AA.
ID PHSL_DESBA
AC P13065;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Periplasmic [NiFeSe] hydrogenase large subunit (EC 1.18.99.1) (NiFeSe
DE hydrogenlyase large chain).
DE Desulfovibrio baculatus (Desulfomicrobium baculatus).
OS Bacteria; Proteobacteria; delta subdivision; Desulfomicrobium.
OX NCBI_TaxID=899;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88058744; PubMed=3316183;

```

RA Menon N.K., Peck H.D. Jr., le Gall J., Przybyla A.E.;  
RT "Cloning and sequencing of the genes encoding the large and small  
RT subunits of the periplasmic (Nifese) hydrogenase of *Desulfovibrio*  
RT *baculatus*."; J. Bacteriol. 169:5401-5407(1987).  
RL [2]  
RN  
RP REVISIONS.  
RA Menon N.K., Peck H.D. Jr., le Gall J., Przybyla A.E.;  
RL J. Bacteriol. 170:4429-4429(1988).  
RN [3]  
RP  
RX X-RAY CRYSTALLOGRAPHY (2.15 ANGSTROMS).  
RA MEDLINE=99306038; PubMed=10378275;  
RA Garcin E., Vernede X., Hatchikian E.C., Volbeda A., Frey M.,  
RA Fontecilla-Camps J.C.;  
RT "The crystal structure of a reduced [Nifese] hydrogenase provides an  
RT image of the activated catalytic center."; J. Bacteriol. 170:4429-4429(1988).  
RL Structure 7:557-566(1999).  
CC -1- CATALYTIC ACTIVITY: 2 reduced ferredoxin + 2 H(+) = 2 oxidized  
CC ferredoxin + H(2).  
CC  
CC -1- COFACTOR: ONE NICKEL ION, TWO 4FE-4S CLUSTERS AND ONE  
CC SELENYCYSTEINE.  
CC  
CC -1- SUBUNIT: HETERODIMER OF A LARGE AND A SMALL SUBUNIT.  
CC -1- SUBCELLULAR LOCATION: Periplasmic.  
CC -1- MISCELLANEOUS: PERHAPS THE LEADER OF THE SMALL SUBUNIT SERVES AS A  
CC TRANSPORT VEHICLE FOR BOTH SUBUNITS.  
CC -1- SIMILARITY: BELONGS TO THE [NIFE]/[NIFSE] HYDROGENASE LARGE  
CC SUBUNIT FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; M18271; AAA23375.2; -  
CC PIR; A33101; HODVLB  
CC  
CC DR PDB; 1CC1; 01-JUN-99.  
CC DR InterPro: IPR001501; NI\_hdl.  
CC DR Pfam; PF00374; Nifese\_Hases; 1.  
CC DR PROSITE; PS00507; NI\_HGENASE\_L1; 1.  
CC DR PROSITE; PS00508; NI\_HGENASE\_L2; 1.  
CC DR Oxidoreductase; Periplasmic; Nickel; Selenium; Selenocysteine;  
CC 3D-structure.  
CC  
CC INIT\_MET 0 0  
CC METAL 51 51 IRON 2.  
CC METAL 70 70 NICKEL.  
CC METAL 73 73 IRON 1 AND NICKEL.  
CC METAL 444 444 IRON 2.  
CC METAL 492 492 NICKEL.  
CC METAL 495 495 IRON 1 AND NICKEL.  
CC METAL 498 498 IRON 2.  
CC SE\_CYS 492 492  
CC  
CC SQ SEQUENCE 513 AA; 56683 MW; AC8285A6F80576FC CRC64;  
CC  
CC Query Match 58.98; Score 33; DB 1; Length 513;  
CC Best Local Similarity 71.4%; Pred. No. 50;  
CC Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
CC  
CC QY 5 PXGMHYS 11  
CC I I I I I  
CC  
CC Db 297 PGLGHYS 303  
CC  
CC RESULT 11  
CC MUTL\_BACSU  
CC ID MUTL\_BACSU STANDARD; PRT; 627 AA.  
CC AC P49850;  
CC DT 01-OCT-1996 (Rel. 34, Created)  
CC DT 01-OCT-1996 (Rel. 34, Last sequence update)  
CC DT 15-JUN-2002 (Rel. 41, Last annotation update)  
CC DE DNA mismatch repair protein mutL.

GN MUTL.  
OS *Bacillus subtilis*.  
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
OX NCBI\_TaxID=1423;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=168;  
RX MEDLINE=96349107; PubMed=8760914;  
RA Ginetti F., Perego M., Albertini A.M., Galizzi A.;  
RT "Bacillus subtilis mutL operon: identification, nucleotide  
RL sequence and mutagenesis."; J. Bacteriol. 170:4429-4429(1988).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=168;  
RX MEDLINE=98044033; PubMed=9384377;  
RA Kunst F., Ogawara N., Moszer I., Albertini A.M., Alloni G.,  
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,  
RA Borriess R., Boursier L., Bruns A., Braun M., Brignell S.C., Bron S.,  
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,  
RA Choi S.K., Codani J.J., Conerton I.F., Cummings N.J., Daniel R.A.,  
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,  
RA Ertian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,  
RA Eritan K.D., Errington J., Fabret C., Ferrari E., Foulger D.,  
RA Chim S.Y., Glaser P., Goffeau A., Gollightly E.J., Grandi G.,  
RA Guisepi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,  
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,  
RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,  
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,  
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,  
RA Lee S.M., Levine A., Liu H., Masuda S., Maue C., Medigue C.,  
RA Medina N., Mellado R.P., Mizuno M., Mostl D., Nakai S., Noback M.,  
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,  
RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,  
RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,  
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadate Y.,  
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scofield F.,  
RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,  
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,  
RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,  
RA Tosato V., Uchiyama S., Vandenberg M., Vannier F., Vassarotti A.,  
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,  
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,  
RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;  
RT "The complete genome sequence of the Gram-positive bacterium *Bacillus*  
RT *subtilis*."; Nature 390:249-256(1997).  
RL Nature 390:249-256(1997).  
CC -1- FUNCTION: THIS PROTEIN IS INVOLVED IN THE REPAIR OF MISMATCHES IN  
CC DNA. IT IS REQUIRED FOR DAM-DEPENDENT METHYL-DIRECTED DNA MISMATCH  
CC REPAIR. MAY ACT AS A "MOLECULAR MATCHMAKER", A PROTEIN THAT  
CC PROMOTES THE FORMATION OF A STABLE COMPLEX BETWEEN TWO OR MORE  
CC DNA-BINDING PROTEINS IN AN ATP-DEPENDENT MANNER WITHOUT ITSELF  
CC BEING PART OF A FINAL EFFECTOR COMPLEX (BY SIMILARITY).  
CC -1- SIMILARITY: BELONGS TO THE DNA MISMATCH REPAIR MUTL/HEX FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; U27343; AAB19236.1; -  
CC EMBL; 299112; CAB13578.1; -  
CC HSP; P23367; 1BKN.  
CC  
CC DR Subtilist; BG11402; mutL.  
CC DR InterPro: IPR003594; ATPbind\_ATPase.  
CC DR InterPro: IPR002099; DNA\_mis\_repair.  
CC DR InterPro: IPR004359; HIS\_KIN\_sig.  
CC DR Pfam; PF01119; DNA\_mis\_repair; 1.  
CC Pfam; PF02518; HATPase-c; 1.  
CC TIGRFAMs; TIGR00585; mutL; 1.

DR PROSITE; PS00058; DNA\_MISMATCH\_REPAIR\_1; 1.  
KW DNA repair; Complete proteome.  
SQ SEQUENCE 627 AA; 70431 MW; 068A0509CC265343 CRC64;

Query Match 58.9%; Score 33; DB 1; Length 627;  
Best Local Similarity 54.5%; Pred. No. 61;  
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 1 EEVVPXGMHYS 11  
| :|||  
Db 488 EMIVLTHYS 498

RESULT 12  
ID KPBA\_CAEEL STANDARD; PRT; 1188 AA.  
AC P34335;  
DT 01-FEB-1994 (Rel. 28, Created)  
DT 15-JUN-2002 (Rel. 41, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Probable phosphorylase B kinase alpha regulatory chain (Phosphorylase  
kinase alpha subunit).  
GN C14B9.8.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;  
OC Rhabditidae; Peleoderinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA STRAIN-Bristol N2;  
RC MEDLINE=94150718; PubMed=7906398;  
RX Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,  
Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,  
Craxton M., Dear S., Du Z., Durbin R., Favell A., Fraser A.,  
Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,  
Johnston L., Jones M., Kershaw J., Kirsten N., Laister N.,  
Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,  
Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Showkeen R.,  
Sims M., Smaldon N., Smith A., Smith K., Sonhammer E., Staden R.,  
Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,  
Waterston R., Watson A., Weinstock L., Wilkinson-Sproat J.,  
Wohlman P.;  
RA "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
elegans";  
RL Nature 368:32-38(1994).  
RN [2]  
RP REVISIONS.  
RA Waterston R.;  
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: PHOSPHORYLASE B KINASE CATALYZES THE PHOSPHORYLATION OF  
SERINE IN CERTAIN SUBSTRATES, INCLUDING TROPONIN I. THE ALPHA  
CHAIN MAY BIND CALMODULIN (BY SIMILARITY).  
CC -1- PATHWAY: Glycogen metabolism.  
CC -1- SIMILARITY: BELONGS TO THE PHOSPHORYLASE B KINASE REGULATORY  
CHAINS FAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; L15188; AAA27954.2; -;  
DR PIR; S44754; S44754.  
DR WormPep; C14B9.8; CE26870.  
KW Hypothetical protein; Glycogen metabolism; Calmodulin-binding.  
SQ SEQUENCE 1188 AA; 135558 MW; DE9BB875F3603863 CRC64;

Query Match 58.9%; Score 33; DB 1; Length 1188;  
Best Local Similarity 54.5%; Pred. No. 1.2e+02;  
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 1 EEVVPXGMHYS 11  
| :|||  
Db 950 EYAEAGIHYS 960

RESULT 13  
VLTF\_BPT5  
ID VLTF\_BPT5 STANDARD; PRT; 1396 AA.  
AC P13390; O48502;  
DT 01-JAN-1990 (Rel. 13, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE L-shaped tail fiber protein (LTF protein).  
GN LTF.  
OS Bacteriophage T5.  
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae;  
OC T5-like viruses.  
OX NCBI\_TaxID=10726;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA MEDLINE=95309401; PubMed=7789514;  
RX Kaliman A.V., Kulshin V.E., Shlyapnikov M.G., Ksenzenko V.N.,  
Kryukov V.M.;  
RA "The nucleotide sequence of the bacteriophage T5 ltf gene";  
RL FEBS Lett. 366:46-48(1995).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Kaliman A.V.;  
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP PRELIMINARY PARTIAL SEQUENCE FROM N.A.  
RX MEDLINE=88289370; PubMed=3267228;  
RA Kaliman A.V., Kryukov V.M., Bayev A.A.;  
RI "The nucleotide sequence of bacteriophage T5 DNA at the region  
between early and late genes";  
RL Nucleic Acids Res. 16:6230-6230(1988).  
CC -1- FUNCTION: NONESSENTIAL PROTEIN THAT MEDIATES BINDING TO THE  
POLYMANNOSE O ANTIGEN.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; X69460; CAA49220.1; -;  
DR EMBL; AJ001191; CAA04591.1; -;  
DR PIR; S01982; S01982.  
KW Late protein.  
FT CONFLICT 986 986 V -> A (IN REF. 2).  
SQ SEQUENCE 1396 AA; 147989 MW; 18CD2192F65FFFC1 CRC64;

Query Match 58.9%; Score 33; DB 1; Length 1396;  
Best Local Similarity 50.0%; Pred. No. 1.4e+02;  
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 2 EEVVPXGMHYS 11  
| :|||  
Db 1360 KTIAGDHYS 1369

RESULT 14  
TKNK\_BOVIN  
ID TKNK\_BOVIN STANDARD; PRT; 126 AA.  
AC P08858;  
DT 01-NOV-1988 (Rel. 09, Created)  
DT 01-NOV-1988 (Rel. 09, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Neurokinin B precursor (NKB) (Neuromedin K).  
GN TAC3 OR NKNB.

OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Bovinae; Bos.  
 OX NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=86313713; PubMed=3462746;  
 RA Kotani H., Hoshimaru M., Nawa H., Nakanishi S.;  
 RT "Structure and gene organization of bovine neuromedin K precursor.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 83:7074-7078(1986).  
 CC -!- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,  
 CC EVOKE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND  
 CC SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH  
 CC MUSCLES.  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 DR EMBL; M14351; AAA30723.1; --  
 DR EMBL; M14347; AAA30723.1; JOINED.  
 DR EMBL; M14348; AAA30723.1; JOINED.  
 DR EMBL; M14349; AAA30723.1; JOINED.  
 DR EMBL; M14350; AAA30723.1; JOINED.  
 DR PIR; A25905; A25905.  
 DR InterPro; IPR003635; Neurokinin.  
 DR InterPro; IPR002040; Tachykinin.  
 DR ProDom; PD020370; Neurokinin; 1.  
 DR PROSITE; PS00267; TACHYKININ; 1.  
 DR Tachykinin; Neuropeptide; Cleavage on pair of basic residues;  
 KW Amidation; Signal.  
 FT SIGNAL 1 20 POTENTIAL.  
 FT PROPEP 21 83  
 FT PEPTIDE 86 95 NEUROKININ B.  
 FT PROPEP 99 126  
 FT MOD\_RES 95 95  
 FT MOD\_RES 95 95  
 SO SEQUENCE 126 AA; 13871 MW; 446EF43498EC059 CRC64;  
 Query Match 57.1%; Score 32; DB 1; Length 126;  
 Best Local Similarity 66.7%; Pred. No. 19;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 Qy 1 EEVVPXGMH 9  
 Db 28 EQVVPGGGH 36  
 I:| | | | |  
 I:| | | | |  
 RESULT 15  
 RR2\_CHLVU  
 ID -RR2\_CHLVU STANDARD; PRT; 267 AA.  
 AC P56351;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE Chloroplast 30S ribosomal protein S2.  
 GN RPS2.  
 OS Chlorella vulgaris.  
 OG Chloroplast.  
 OC Eukaryota; Viridiplantae; Chlorophyta; Trebouxiophyceae; Chlorellales;  
 OC Chlorellaceae; Chlorella.  
 OX NCBI\_TaxID=3077;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-TAM C-27 / Tamiya;  
 RX MEDLINE=97303241; PubMed=9159184;  
 RA Wakasugi T., Nagai T., Kapoor M., Sugita M., Ito S.,

TSudzuki J., Nakashima K., Tsudzuki T., Suzuki Y., Hamada A., Ohta T.,  
 Inamura A., Yoshinaga K., Sugitara M.;  
 RT "Complete nucleotide sequence of the chloroplast genome from the  
 green alga Chlorella vulgaris: the existence of genes possibly  
 involved in chloroplast division.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 94:5967-5972(1997).  
 CC -!- SIMILARITY: BELONGS TO THE S2P FAMILY OF RIBOSOMAL PROTEINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 DR EMBL; AB001684; BAA57862.1; --  
 DR InterPro; IPR001865; Ribosomal\_S2.  
 DR Pfam; PF00318; Ribosomal\_S2; 1.  
 DR PRINTS; PR00395; RIBOSOMALS2.  
 DR TIGFAMS; TIGR01011; rpsB\_bact; 1.  
 DR PROSITE; PS00962; RIBOSOMAL\_S2\_1; FALSE\_NEG.  
 DR PROSITE; PS00963; RIBOSOMAL\_S2\_2; 1.  
 KW Ribosomal protein; Chloroplast.  
 SQ SEQUENCE 267 AA; 30699 MW; 7903075340BD900F CRC64;  
 Query Match 57.1%; Score 32; DB 1; Length 267;  
 Best Local Similarity 40.0%; Pred. No. 40;  
 Matches 4; Conservative 4; Mismatches 2; Indels 0; Gaps 0;  
 Qy 1 EEVVPXGMHY 10  
 Db 8 EDMIQSGMHF 17  
 I:| | | | |  
 I:| | | | |  
 RESULT 16  
 CGD2\_RAT  
 ID -CGD2\_RAT STANDARD; PRT; 288 AA.  
 AC Q04827;  
 DT 01-OCT-1993 (Rel. 27, Created)  
 DT 01-OCT-1993 (Rel. 27, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE G1/S-specific cyclin D2 (Vin-1 proto-oncogene).  
 GN CCND2 OR VIN-1.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93275661; PubMed=8502486;  
 RA Hanna Z., Jankowski M., Tremblay P., Jiang X.M., Milatovich A.,  
 Francke U., Jolicœur P.;  
 RT "The Vin-1 gene, identified by provirus insertional mutagenesis, is  
 the cyclin D2.";  
 RL Oncogene 8:1661-1666(1993).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95011623; PubMed=7926809;  
 RA Hosokawa Y., Onga T., Nakashima K.;  
 RT "Induction of D2 and D3 cyclin-encoding genes during promotion of the  
 G1/S transition by prolactin in rat Nb2 cells.";  
 RL Gene 147:249-252(1994).  
 CC -!- FUNCTION: ESSENTIAL FOR THE CONTROL OF THE CELL CYCLE AT THE G1/S  
 CC (START) TRANSITION.  
 CC -!- SUBUNIT: INTERACTS WITH THE CDK4 AND CDK6 PROTEIN KINASES TO FORM  
 CC A SERINE/THREONINE KINASE HOLOENZYME COMPLEX. THE CYCLIN SUBUNIT  
 CC IMPARTS SUBSTRATE SPECIFICITY TO THE COMPLEX.  
 CC -!- SIMILARITY: BELONGS TO THE CYCLIN FAMILY. CYCLIN D SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its



```

RESULT 17
CGD2_HUMAN STANDARD; PRT; 289 AA.
ID CGD2_HUMAN STANDARD; PRT; 289 AA.
P30279; Q13955;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Gl/S-specific cyclin D2.
OS CCND2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A.
MEDLINE=92347851; PubMed=1386336;
RX Xiong Y., Menninger J., Beach D., Ward D.C.;
RA "Molecular cloning and chromosomal mapping of CCND genes encoding
RT human D-type cyclins.";
RL Genomics 13:575-584(1992).
RN [2]
SEQUENCE FROM N.A.
RX MEDLINE=93205384; PubMed=8455931;
RA Palmero I., Holder A., Sinclair A.J., Dickson C., Peters G.;
RT "Cyclins D1 and D2 are differentially expressed in human B-lymphoid
RL cell lines.";
RN [3]
Oncogene 8:1049-1054(1993).
RN [4]
SEQUENCE FROM N.A.
RX Miyajima N.;
RA Submitted (MAR-1993) to the EMBL/GenBank/DBJ databases.
RN [5]
SEQUENCE FROM N.A.
RX Tissue=Bone marrow;
RA Strausberg R.;
RN Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
RN [6]
SEQUENCE OF 1-240 FROM N.A.
RX Tissue=Placenta;
RA MEDLINE=92347850; PubMed=1386335;
RN Inaba T., Matsushime H., Roussel M.F., Sherr C.J.,
RA Look A.T.;
RX "Genomic organization, chromosomal localization, and independent
RT expression of human cyclin D genes.";
RL Genomics 13:565-574(1992).

```

RESULT	18
CGD2_MOUSE	
ID	CGD2_MOUSE STANDARD; PRT; 289 AA..
AC	P30280;
DT	01-APR-1993 (Rel. 25, Created)
DT	01-APR-1993 (Rel. 25, Last sequence update)
DT	16-OCT-2001 (Rel. 40, Last annotation update)
DE	G1/S-specific cyclin D2.
DE	CCND2 OR CYL-2.
GN	Mus musculus (Mouse).
OS	Mus musculus (Mouse).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX	NCBI_TaxID=10090;
RN	[1]
RN	SEQUENCE FROM N.A.
RX	MEDLINE=92196134; PubMed=1372445;
RA	Kiyokawa H., Busquets X., Powell C.T., Ngo L., Rifkind R.A.,
RA	Marks P.A.;
RT	"Cloning of a D-type cyclin from murine erythro leukemia cells.";
RL	C'ong. Natl. Acad. Sci. U.S.A. 89:2444-2447(1992).
RN	[2]
RN	SEQUENCE FROM N.A.
RP	MEDLINE=91235305; PubMed=1827757;
RX	Matsushima H., Roussel M.F., Ashmun R.A., Sherr C.J.;
RA	Matsushima H., Roussel M.F., Ashmun R.A., Sherr C.J.;
RT	"Colony-stimulating factor 1 regulates novel cyclins during the G1
RT	phase of the cell cycle.";

Cell 65:701-713(1991).

-1- FUNCTION: ESSENTIAL FOR THE CONTROL OF THE CELL CYCLE AT THE G1/S (START) TRANSITION.

-1- SUBUNIT: INTERACTS WITH THE CDK4 AND CDK6 PROTEIN KINASES TO FORM A SERINE/THREONINE KINASE HOLOENZYME COMPLEX. THE CYCLIN SUBUNIT IMPARTS SUBSTRATE SPECIFICITY TO THE COMPLEX.

-1- SIMILARITY: BELONGS TO THE CYCLIN FAMILY. CYCLIN D SUBFAMILY.

-----

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

EMBL: M83749; AAA37519.1; -  
 EMBL: M86182; AAA37503.1; -  
 PIR: B40035; B40035.  
 PIR: A41984; A41984.  
 MGI: 88314; Cnd2.  
 InterPro: IPR004366; Cyclin.  
 InterPro: IPR004367; Cyclin\_Cterm.  
 Pfam: PF00134; cyclin; 1.  
 Pfam: PF02984; cyclin\_C; 1.  
 SMART: SM00385; CYCLIN; 1.  
 PROSITE: PS00292; CYCLINS; 1.  
 KW Cyclin; Cell cycle; Cell division; Multigene family.  
 SQ SEQUENCE 289 AA; 32897 MW; 58F322771DD1DA3D CRG64;

Query Match 57.1%; Score 32; DB 1; Length 289;  
 Best Local Similarity 60.0%; Pred. No. 44;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMHY 10  
 ||| | | |  
 Db 73 EEVFPPLAMNY 82

RESULT 19

CGDI\_BRARE STANDARD; PRT; 291 AA.

AC Q90459;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE G1/S-specific cyclin D1.  
 GN CYCD1  
 OS Brachydanio rerio (Zebrafish) (Danio rerio).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
 OC Cyprinidae; Danio.  
 OX NCBI\_TaxID=7955;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96138542; PubMed=8547308;  
 RA Yarden A., Salomon D., Geiger B.;  
 RT "Zebrafish cyclin D1 is differentially expressed during early embryogenesis";  
 RL Biochim. Biophys. Acta 1264:257-260(1995).  
 CC -1- FUNCTION: ESSENTIAL FOR THE CONTROL OF THE CELL CYCLE AT THE G1/S (START) TRANSITION.  
 CC -1- SUBUNIT: INTERACTS WITH THE CDK4 AND CDK6 PROTEIN KINASES TO FORM A SERINE/THREONINE KINASE HOLOENZYME COMPLEX. THE CYCLIN SUBUNIT IMPARTS SUBSTRATE SPECIFICITY TO THE COMPLEX.  
 CC -1- SIMILARITY: BELONGS TO THE CYCLIN FAMILY. CYCLIN D SUBFAMILY.

-----

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

EMBL: M83749; AAA37519.1; -  
 EMBL: M86182; AAA37503.1; -  
 PIR: B40035; B40035.  
 PIR: A41984; A41984.  
 MGI: 88314; Cnd2.  
 InterPro: IPR004366; Cyclin.  
 InterPro: IPR004367; Cyclin\_Cterm.  
 Pfam: PF00134; cyclin; 1.  
 Pfam: PF02984; cyclin\_C; 1.  
 SMART: SM00385; CYCLIN; 1.  
 PROSITE: PS00292; CYCLINS; 1.  
 KW Cyclin; Cell cycle; Cell division; Multigene family.  
 SQ SEQUENCE 289 AA; 32897 MW; 58F322771DD1DA3D CRG64;

Query Match 57.1%; Score 32; DB 1; Length 289;  
 Best Local Similarity 60.0%; Pred. No. 44;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMHY 10  
 ||| | | |  
 Db 73 EEVFPPLAMNY 82

or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL: X87581; CAA60885.1; -  
 ZFIN: ZDB-GENE-980526-176; cycdl.  
 InterPro: IPR004366; Cyclin.  
 InterPro: IPR004367; Cyclin\_Cterm.  
 Pfam: PF00134; cyclin; 1.  
 Pfam: PF02984; cyclin\_C; 1.  
 SMART: SM00385; CYCLIN; 1.  
 PROSITE: PS00292; CYCLINS; 1.  
 KW Cyclin; Cell cycle; Cell division.  
 SQ SEQUENCE 291 AA; 33067 MW; FA5274CB1B46D5EF CRG64;

Query Match 57.1%; Score 32; DB 1; Length 291;  
 Best Local Similarity 60.0%; Pred. No. 44;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMHY 10  
 ||| | | |  
 Db 75 EEVFPPLAMNY 84

# RESULT 20

CGDI\_XENLA STANDARD; PRT; 291 AA.

AC P50755;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE G1/S-specific cyclin D1.  
 GN CCND1.  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipiloidea; Pipidae;  
 OC Xenopodinae; Xenopus.  
 OX NCBI\_TaxID=8355;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Cockerill M.J., Hunt T.;  
 RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.

-1- FUNCTION: ESSENTIAL FOR THE CONTROL OF THE CELL CYCLE AT THE G1/S (START) TRANSITION.  
 CC -1- SUBUNIT: INTERACTS WITH THE CDK4 AND CDK6 PROTEIN KINASES TO FORM A SERINE/THREONINE KINASE HOLOENZYME COMPLEX. THE CYCLIN SUBUNIT IMPARTS SUBSTRATE SPECIFICITY TO THE COMPLEX.  
 CC -1- SIMILARITY: BELONGS TO THE CYCLIN FAMILY. CYCLIN D SUBFAMILY.

-----

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

EMBL: X89475; CAA61664.1; -  
 InterPro: IPR004366; Cyclin.  
 InterPro: IPR004367; Cyclin\_Cterm.  
 Pfam: PF00134; cyclin; 1.  
 Pfam: PF02984; cyclin\_C; 1.  
 SMART: SM00385; CYCLIN; 1.  
 PROSITE: PS00292; CYCLINS; 1.  
 KW Cyclin; Cell cycle; Cell division; Multigene family.  
 SQ SEQUENCE 291 AA; 32953 MW; A4747CSBD1679087 CRG64;

Query Match 57.1%; Score 32; DB 1; Length 291;  
 Best Local Similarity 60.0%; Pred. No. 44;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMHY 10  
 ||| | | |  
 Db 73 EEVFPPLAMNY 82

```

RESULT 21
CGD2_CHICK STANDARD; PRT; 291 AA.
AC P49706;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE GL/S-specific cyclin D2.
GN CCND2.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE=96144302; PubMed=8566807;
RA Li H., Grenet J., Kidd V.J.;
RT "Structure and gene expression of avian cyclin D2.";
RL Gene 167:341-342(1995).
CC -1- FUNCTION: ESSENTIAL FOR THE CONTROL OF THE CELL CYCLE AT THE G1/S
CC (START) TRANSITION.
CC -1- SUBUNIT: INTERACTS WITH THE CDK4 AND CDK6 PROTEIN KINASES TO FORM
CC A SERINE/THREONINE KINASE HOLOENZYME COMPLEX. THE CYCLIN SUBUNIT
CC IMPARTS SUBSTRATE SPECIFICITY TO THE COMPLEX.
CC -1- SIMILARITY: BELONGS TO THE CYCLIN FAMILY. CYCLIN D SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X89476; CAA61665.1; -
CC EMBL; X83503; CAA58493.1; -
CC InterPro; IPR004366; Cyclin.
CC InterPro; IPR004367; Cyclin_Cterm.
CC Pfam; PF00134; cyclin; 1.
CC Pfam; PF02984; cyclin_C; 1.
CC SMART; SM00385; CYCLIN; 1.
CC PROSITE; PS00292; CYCLINS; 1.
CC Cyclin; Cell cycle; Cell division; Multigene family.
KW CYCLIN; Cell cycle; Cell division; Multigene family.
SQ SEQUENCE 291 AA; 32959 MW; 9A290F04F1531E89 CRC64;

Query Match 57.1%; Score 32; DB 1; Length 291;
Best Local Similarity 60.0%; Pred. No. 44;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMHY 10
DB ||| | | |
74 EEVFPAMNY 83

RESULT 22
CGD2_XENLA STANDARD; PRT; 291 AA.
AC P53782;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE GL/S-specific cyclin D2.
GN CCND2.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
SEQUENCE FROM N.A.
RX Cockerill M.J., Hunt T.;
RA Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.

```

DR EMBL: U40844; AAA83271.1; -  
 DR InterPro: IPR004366; Cyclin.  
 DR InterPro: IPR004367; Cyclin\_Cterm.  
 DR Pfam: PF00134; cyclin; 1.  
 DR PFam: PF02984; cyclin\_C; 1.  
 DR SMART: SM00385; CYCLIN; 1.  
 DR PROSITE: PS00292; CYCLINS; 1.  
 DR Cyclin: Cell cycle; Cell division; Multigene family.  
 KW CYCLIN; Cell cycle; Cell division; Multigene family.  
 SQ SEQUENCE 292 AA; 3263 MW; 7B543029DB45A67D CRC64;  
 Query Match 57.1%; Score 32; DB 1; Length 292;  
 Best Local Similarity 60.0%; Pred. No. 44;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 EEVVPXGMHY 10  
 DB ||| | | | |  
 75 EEVFPPLAMNY 84  
 RESULT 24  
 CGD3\_HUMAN STANDARD; PRT; 292 AA.  
 AC P30281; Q96F49;  
 DT 01-APR-1993 (Rel. 25, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE G1/S-specific cyclin D3.  
 GN CCND3.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-92347851; PubMed-1386336;  
 RA Xiong Y., Menninger J., Beach D., Ward D.C.;  
 RT "Molecular cloning and chromosomal mapping of CCND genes encoding  
 human D-type cyclins.";  
 RL Genomics 13:575-584(1992).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-93015922; PubMed-1383201;  
 RA Motokura T., Keyomarsi K., Kronenberg H.M., Arnold A.;  
 RT "Cloning and characterization of human cyclin D3, a cDNA closely  
 related in sequence to the PRAD1/cyclin D1 proto-oncogene.";  
 RL J. Biol. Chem. 267:20412-20415(1992).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX Strausberg R.;  
 RA Tissue-Lung;  
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE OF 52-237 FROM N.A.  
 RX Tissue-Placenta;  
 RA Inaba T., Matsushime H., Valentine M., Roussel M.F., Sherr C.J.,  
 RA Look A.T.;  
 RT "Genomic organization, chromosomal localization, and independent  
 expression of human cyclin D genes.";  
 RL Genomics 13:565-574(1992).  
 CC -1- FUNCTION: ESSENTIAL FOR THE CONTROL OF THE CELL CYCLE AT THE G1/S  
 CC -1- SUBUNIT: INTERACTS WITH THE CDK4 AND CDK6 PROTEIN KINASES TO FORM  
 CC A SERINE/THREONINE KINASE HOLOENZYME COMPLEX. THE CYCLIN SUBUNIT  
 CC IMPARTS SUBSTRATE SPECIFICITY TO THE COMPLEX.  
 CC -1- SIMILARITY: BELONGS TO THE CYCLIN FAMILY. CYCLIN D SUBFAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>)

or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).  
 CC  
 CC EMBL: M90814; AAA51927.1; -  
 DR EMBL: M92287; AAA52137.1; -  
 DR EMBL: BC011616; AAH11616.1; -  
 DR EMBL: M88087; AAA51929.1; -  
 DR EMBL: M88084; AAA51929.1; JOINED.  
 DR EMBL: M88085; AAA51929.1; JOINED.  
 DR EMBL: M88086; AAA51929.1; JOINED.  
 DR PIR: B42822; B42822.  
 DR PIR: A44022; A44022.  
 DR GENE: HGNC:1585; CCND3.  
 DR MIM: 123834; -  
 DR InterPro: IPR004366; Cyclin.  
 DR InterPro: IPR004367; Cyclin\_Cterm.  
 DR Pfam: PF00134; cyclin; 1.  
 DR Pfam: PF02984; cyclin\_C; 1.  
 DR SMART: SM00385; CYCLIN; 1.  
 DR PROSITE: PS00292; CYCLINS; 1.  
 DR Cyclin: Cell cycle; Cell division; Multigene family.  
 KW CYCLIN; Cell cycle; Cell division; Multigene family.  
 FT CONFLICT 259 259 S -> A (IN REF. 1).  
 SQ SEQUENCE 292 AA; 32519 MW; 1657B1604FEB0029 CRC64;  
 Query Match 57.1%; Score 32; DB 1; Length 292;  
 Best Local Similarity 60.0%; Pred. No. 44;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 EEVVPXGMHY 10  
 DB ||| | | | |  
 75 EEVFPPLAMNY 84  
 RESULT 25  
 CGD1\_HUMAN STANDARD; PRT; 295 AA.  
 ID C24385;  
 AC P24385; (Rel. 21, Created)  
 DT 01-MAR-1992 (Rel. 21, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE G1/S-specific cyclin D1 (PRAD1 oncogene) (BCL-1 oncogene).  
 GN CCND1 OR PRAD1 OR BCL1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-91194766; PubMed-1826542;  
 RA Motokura T., Bloom T., Kim H.G., Jueppner H., Ruderman J.V.,  
 RA Kronenberg H.M., Arnold A.;  
 RT "A novel cyclin encoded by a bcl1-linked candidate oncogene.";  
 RL Nature 350:512-515(1991).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-92005671; PubMed-1833066;  
 RA Lew D.J., Dulic V., Reed S.I.;  
 RT "Isolation of three novel human cyclins by rescue of G1 cyclin (Cln)  
 function in yeast.";  
 RL Cell 66:1197-1206(1991).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-91235304; PubMed-1827756;  
 RA Xiong Y., Connolly T., Fletcher B., Beach D.;  
 RT "Human D-type cyclin.";  
 RL Cell 65:691-699(1991).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-92017758; PubMed-1833629;  
 RA Withers D.A., Harvey R.C., Faust J.B., Melnyk O., Carey K.,  
 RA Meeker T.C.;  
 RT "Characterization of a candidate bcl-1 gene.";  
 RL Mol. Cell. Biol. 11:4846-4853(1991).  
 RN [5]

RP SEQUENCE FROM N.A.  
RX MEDLINE=94264323; PubMed=8204893;  
RA Rimokh R., Berger F., Bastard C., Klein B., French M., Archimbaud E.,  
EA Rouault J.-P., Santa Lucia B., Duret L., Vuillaume M.;  
RT "Rearrangement of CCND1 (BCL1/PRAD1) 3' untranslated region in  
RL mantle-cell lymphomas and t(11q13)-associated leukemias.";  
RN Blood 83:3689-3696(1994).  
[6]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RA Strausberg R.;  
RN Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
[7]  
RP INTERACTION WITH CDK4 AND CDK6.  
RX MEDLINE=94134440; PubMed=8302605;  
RA Bates S., Bonetta L., McAllian D., Parry D., Holder A., Dickson C.,  
RA Peters G.;  
RT "CDK6 (P13TIRE) and CDK4 (P3K-J3) are a distinct subset of the  
RL cyclin-dependent kinases that associate with cyclin D1.";  
RN Oncogene 9:71-79(1994).  
CC -!- FUNCTION: ESSENTIAL FOR THE CONTROL OF THE CELL CYCLE AT THE G1/S  
CC (START) TRANSITION.  
CC -!- SUBUNIT: INTERACTS WITH THE CDK4 AND CDK6 PROTEIN KINASES TO FORM  
CC A SERINE/THREONINE KINASE HOLOENZYME COMPLEX. THE CYCLIN SUBUNIT  
CC IMPARTS SUBSTRATE SPECIFICITY TO THE COMPLEX.  
CC -!- DISEASE: INVOLVED IN B-LYMPHOCTIC MALIGNANCY (PARTICULARLY  
CC MANTLE-CELL LYMPHOMA (MCL)) BY A CHROMOSOMAL TRANSLOCATION  
CC T(11;14)(q13;q32) THAT INVOLVES CCND1 AND IMMUNOGLOBULIN GENE  
CC REGIONS (BCL1 ONCOGENE). ACTIVATION OF CCND1 MAY BE ONCOGENIC BY  
CC DIRECTLY ALTERING PROGRESSION THROUGH THE CELL CYCLE.  
CC -!- DISEASE: INVOLVED IN A SUBSET OF PARATHYROID ADENOMAS BY A  
CC CHROMOSOMAL TRANSLOCATION T(11;11)(q13;p15) THAT INVOLVES CCND1  
CC AND THE PARATHYROID HORMONE (PTH) ENHANCER (PRAD1 ONCOGENE).  
CC -!- SIMILARITY: BELONGS TO THE CYCLIN FAMILY. CYCLIN D SUBFAMILY.  
CC -!- DATABASE: NAME-Atlas Genet. Cytogenet. Oncol. Haematol.;  
CC WWW="http://www.infobiogen.fr/services/chronocancer/Genes/BCL1.html".  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL; X59798; CAA42470.1; -;  
DR EMBL; M74092; -; NOT\_ANNOTATED\_CDS.  
DR EMBL; M64349; AAA52136.1; -;  
DR EMBL; M73554; AAA58392.1; -;  
DR EMBL; Z23022; CAA80558.1; -;  
DR EMBL; BC000076; AAH00076.1; -;  
DR EMBL; BC001501; AAH01501.1; -;  
DR EMBL; BC014078; AAH14078.1; -;  
DR PIR; B40268; B40268.  
DR PIR; S14794; S14794.  
DR PIR; A41523; A41523.  
DR PIR; A40034; A40034.  
DR PIR; S34295; S34295.  
DR Genew; HGNC:1582; CCND1.  
DR MIM; 151400; -;  
DR MIM; 168461; -;  
DR InterPro; IPR004366; Cyclin.  
DR InterPro; IPR004367; Cyclin\_Cterm.  
DR Pfam; PF00134; cyclin; 1.  
DR Pfam; PF02984; cyclin\_C; 1.  
DR SMART; SM00385; CYCLIN; 1.  
DR PROSITE; PS00292; CYCLINS; 1.  
KW Cyclin; Cell cycle; Cell division; Multigene family; Proto-oncogene;  
KW Chromosomal translocation.  
FT CONFLICT 130 130 N -> G (IN REF. 3).  
FT CONFLICT 168 169 MP -> IA (IN REF. 2).  
FT CONFLICT 188 188 L -> S (IN REF. 3).  
SQ SEQUENCE 295 AA; 33729 MW; 3CC00C9905F58D3A CRC64;

Query Match 57.1%; Score 32; DB 1; Length 295;  
Best Local Similarity 60.0%; Pred. No. 45;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
QY 1 EEVVPXGMHY 10  
| | | | | | | | | |  
DB 75 EEVFPPLAMNY 84

Search completed: June 10, 2003, 13:40:18  
Job time : 5.5 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2003, 13:30:25 ; Search time 25.7857 seconds  
(without alignments)  
87.898 Million cell updates/sec

Title: US-09-909-164-7  
Perfect score: 56  
Sequence: 1 EEVVPXGMHYS 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues  
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 75 summaries

Database : SPTREMBL\_21.\*

- 1: sp\_archaea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mhc.\*
- 8: sp\_organelle.\*
- 9: sp\_phase.\*
- 10: sp\_plant.\*
- 11: sp\_rodent.\*
- 12: sp\_virus.\*
- 13: sp\_vertebrate.\*
- 14: sp\_unclassified.\*
- 15: sp\_rvirus.\*
- 16: sp\_bacteriap.\*
- 17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	38	67.9	1063	Q8RG86	Q8rg86 fusobacteri
2	37	66.1	308	Q9X2E2	Q9x2e2 thermotoga
3	37	66.1	322	Q9HLH8	Q9hlh8 thermoplasm
4	37	66.1	1057	Q99UR5	Q99ur5 staphylococ
5	36	64.3	208	Q46486	Q46486 corynebacte
6	36	64.3	252	Q8VWP1	Q8vwp1 anabaena sp
7	36	64.3	819	Q9AVK4	Q9avk4 pisum sativ
8	35	62.5	139	Q57489	Q57489 bacteroides
9	35	62.5	156	Q9PC35	Q9pc35 xylella fas
10	35	62.5	233	Q40479	Q40479 nicotiana t
11	35	62.5	237	Q9LW50	Q9lw50 nicotiana s
12	35	62.5	317	Q38317	Q38317 lactobacill
13	35	62.5	425	Q9XVK4	Q9xvk4 caenorhabd
14	35	62.5	510	Q9SA71	Q9sa71 arabidopsis
15	35	62.5	2042	Q8TZ07	Q8tz07 methanopyru
16	34	60.7	264	Q27902	Q27902 methanobact

17	34	60.7	279	16	Q9RXN9	Q9rxn9 deinococcus
18	34	60.7	285	16	Q98HU6	Q98hu6 rhizobium l
19	34	60.7	350	16	Q9RW92	Q9rw92 deinococcus
20	34	60.7	355	16	Q9XAM3	Q9xam3 streptomyce
21	34	60.7	360	17	Q27679	Q27679 methanobact
22	34	60.7	425	16	Q9KBA1	Q9kba1 bacillus ha
23	34	60.7	495	5	O16912	O16912 caenorhabdi
24	34	60.7	637	5	Q9Y0Y6	Q9y0y6 drosophila
25	34	60.7	678	12	Q9ELX6	Q9elx6 cercopithe
26	34	60.7	748	12	Q9YR01	Q9yr01 ranid herpe
27	34	60.7	1028	16	Q8YJ11	Q8yj11 brucella me
28	34	60.7	2438	5	Q9VQL7	Q9vql7 drosophila
29	33	58.9	156	3	Q12479	Q12479 saccharomy
30	33	58.9	216	16	Q9RDC1	Q9rdcl streptomyce
31	33	58.9	257	8	Q99011	Q99011 prototheca
32	33	58.9	273	4	Q96N44	Q96n44 homo sapien
33	33	58.9	306	16	Q8XVQ7	Q8xvq7 raistonias
34	33	58.9	367	16	Q916J1	Q916j1 pseudomonas
35	33	58.9	387	16	Q98FX1	Q98fx1 rhizobium l
36	33	58.9	426	3	Q9HG99	Q9hg99 mortierella
37	33	58.9	441	16	Q9KT98	Q9kt98 vibrio chol
38	33	58.9	466	16	Q9PKH7	Q9pkh7 chlamydia m
39	33	58.9	466	16	Q84218	Q84218 chlamydia t
40	33	58.9	534	17	Q29966	Q29966 archaeoglob
41	33	58.9	542	12	Q9WCW0	Q9wcw0 avian infec
42	33	58.9	545	10	Q9T025	Q9t025 arabidopsis
43	33	58.9	556	12	Q9QGT4	Q9qgt4 avian infec
44	33	58.9	558	12	Q9QGT3	Q9qgt3 avian infec
45	33	58.9	564	12	Q9QGT1	Q9qgt1 avian infec
46	33	58.9	565	12	Q9QGT5	Q9qgt5 avian infec
47	33	58.9	565	12	Q9QGT2	Q9qgt2 avian infec
48	33	58.9	567	5	Q9VEG2	Q9veg2 drosophila
49	33	58.9	572	4	Q8TEP5	Q8tep5 homo sapien
50	33	58.9	601	16	Q984F4	Q984f4 rhizobium l
51	33	58.9	716	11	P70521	P70521 rattus norv
52	33	58.9	796	10	Q8VY35	Q8vy35 zea mays (m
53	33	58.9	1442	17	Q96YH5	Q96yh5 sulfolobus
54	33	58.9	1471	4	Q8TEN9	Q8ten9 homo sapien
55	33	58.9	3472	1	O74056	O74056 cenarchaeum
56	32	57.1	103	11	Q9D0H9	Q9d0h9 mus musculu
57	32	57.1	153	5	O76217	O76217 anopheles g
58	32	57.1	153	13	P79919	P79919 xenopus lae
59	32	57.1	156	11	Q9D8L9	Q9d8l9 mus musculu
60	32	57.1	190	13	O57481	O57481 stizostedio
61	32	57.1	191	11	Q99NB4	Q99nb4 rattus norv
62	32	57.1	197	16	Q9ZCF9	Q9zcf9 rickettsia
63	32	57.1	225	10	Q40129	Q40129 lycopersico
64	32	57.1	233	16	Q97166	Q97166 clostridium
65	32	57.1	235	5	Q9V5M0	Q9v5m0 drosophila
66	32	57.1	240	11	Q9DB09	Q9db09 mus musculu
67	32	57.1	270	16	Q9R9Q3	Q9r9q3 rhizobium m
68	32	57.1	291	13	Q8QFP4	Q8qfp4 brachydanio
69	32	57.1	322	17	Q978P6	Q978p6 thermoplasm
70	32	57.1	329	5	Q8SVQ5	Q8svq5 encephalito
71	32	57.1	335	5	O01456	O01456 caenorhabdi
72	32	57.1	341	2	O44548	O44548 azotobacter
73	32	57.1	363	16	Q913T4	Q913t4 pseudomonas
74	32	57.1	370	2	Q9AGZ9	Q9agz9 pseudomonas
75	32	57.1	374	17	O27916	O27916 methanobact

ALIGNMENTS

RESULT 1				
Q8RG86	Q8RG86	PRELIMINARY;	PRT;	1063 AA.
ID	Q8RG86			
AC	Q8RG86			
DT	01-JUN-2002 (TREMBlrel. 21, Created)			
DT	01-JUN-2002 (TREMBlrel. 21, Last sequence update)			
DT	01-JUN-2002 (TREMBlrel. 21, Last annotation update)			
DE	Carbamoyl-phosphate synthase large chain (EC 6.3.5.5).			
GN	FN0422.			

```

OS Fusobacterium nucleatum (subsp. nucleatum).
OC Bacteria; Fusobacteria; Fusobacterium.
OX NCBI_TaxID=76856;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 25586;
RX MEDLINE=21886394; PubMed=11889109;
RA Kapatal V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A.,
RA Bhattacharya A., Bartman A., Gardner W., Grechkin G., Zhu L.,
RA Vasieva O., Chu L., Kogan Y., Chaga O., Goltsman E., Bernal A.,
RA Larsen N., D'Souza M., Walunas T., Pusch G., Haselkorn R.,
RA Fomstein M., Kyripides N., Overbeek R.;
RT "Genome sequence and analysis of the oral bacterium Fusobacterium
RT nucleatum strain ATCC 25586."
RL J. Bacteriol. 184:2005-2018(2002).
RW EMBL; AE010554; AAL94625.1; -
KW Ligase; Complete proteome.
SQ SEQUENCE 1063 AA; 118008 MW; 39700E10B7CCB411 CRC64;

Query Match 67.9%; Score 38; DB 16; Length 1063;
Best Local Similarity 60.0%; Pred. No. 48;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 EVVPXGMHYS 11
Db 195 EIVPGLNYS 204
||| |::||

RESULT 2
ID Q9X2E2 PRELIMINARY; PRT; 308 AA.
AC Q9X2E2;
DT 01-NOV-1999 (TReMBLrel. 12, Created)
DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE FTSH protease activity modulator HFLK.
GN TM1822.
OS Thermotoga maritima.
OC Bacteria; Thermotogae; Thermotogales; Thermotogaceae; Thermotoga.
OX NCBI_TaxID=2336;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MSB8 / DSM 3109;
RX MEDLINE=99287316; PubMed=10360571;
RA Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
RA Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
RA McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
RA Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
RA Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
RT "Evidence for lateral gene transfer between Archaea and Bacteria from
RT genome sequence of Thermotoga maritima."
RL Nature 399:323-329(1999).
RW EMBL; AE001819; AAD36885.1; -
RX TIGR; TM1822.
DR InterPro; IPR001107; Band 7.
DR InterPro; IPR003130; GED.
DR InterPro; IPR001972; Stomatin.
DR Pfam; PF01145; Band 7; 1.
DR PRINTS; PR00721; STOMATIN.
DR SMART; SM00302; GED; 1.
DR SMART; SM00244; PHB; 1.
KW Protease; Complete proteome.
SQ SEQUENCE 308 AA; 34778 MW; ADE03603E5101A9D CRC64;

Query Match 66.1%; Score 37; DB 16; Length 308;
Best Local Similarity 75.0%; Pred. No. 21;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMHY 10
Db 41 VVPXGIHY 48
||| |::||

```

```

RESULT 3
Q9HLH8 PRELIMINARY; PRT; 322 AA.
ID Q9HLH8;
AC Q9HLH8;
DT 01-VAR-2001 (TReMBLrel. 16, Created)
DT 01-VAR-2001 (TReMBLrel. 16, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE Glucose-fructose oxidoreductase related protein.
GN TA0250.
OS Thermoplasma acidophilum.
OC Archaea; Euryarchaeota; Thermoplasmata; Thermoplasmatales;
OC Thermoplasmataceae; Thermoplasma.
OX NCBI_TaxID=2303;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DSM 1728;
RX MEDLINE=20479972; PubMed=11029001;
RA Ruepp A., Graml W., Santos-Martinez M.-L., Koretke K.K., Volker C.,
RA Mewes H.-W., Frishman D., Stocker S., Lupas A.N., Baumeister W.;
RT "The genome sequence of the thermoacidophilic scavenger Thermoplasma
RT acidophilum."
RL Nature 407:508-513(2000).
DR EMBL; AL445063; CAC11395.1; -
DR InterPro; IPR000683; GFO_IDH_MoCA.
DR Pfam; PF01408; GFO_IDH_MoCA; 1.
KW Complete proteome.
SQ SEQUENCE 322 AA; 36918 MW; B8C239E71009D167 CRC64;

Query Match 66.1%; Score 37; DB 17; Length 322;
Best Local Similarity 75.0%; Pred. No. 22;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMHY 10
Db 66 VVPDGLHY 73
||| |::||

RESULT 4
Q99UR5 PRELIMINARY; PRT; 1057 AA.
ID Q99UR5;
AC Q99UR5;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT 01-VAR-2002 (TReMBLrel. 20, Last annotation update)
DE Carbamoyl-phosphate synthase large chain.
GN PYRAB OR SAV1203 OR SA1046.
OS Staphylococcus aureus (strain Mu50 / ATCC 700699), and
OS Staphylococcus aureus (strain N315).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=158878; 158879;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=S.aureus (strain Mu50), and S.aureus (strain N315);
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Hiramatsu K.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus."
RL Lancet 357:1225-1240(2001).
RW EMBL; AP003361; BAB57365.1; -
DR EMBL; AP003132; BAB42298.1; -
DR HSSP; P00968; IC50.
DR InterPro; IPR000901; CPSase.
DR InterPro; IPR004362; MGS_like.
DR InterPro; IPR000169; SHprot_acsite.
DR Pfam; PF00289; CPSase_L_chain; 2.

```



DR pfam: PF02786; CPSase\_L\_D2; 2.  
 DR pfam: PF02787; CPSase\_L\_D3; 1.  
 DR pfam: PF02142; MGS; 1.  
 DR PRINTS; PRO0098; CPSASE.  
 DR PROSITE; PS00866; CPSASE\_1; 2.  
 DR PROSITE; PS00867; CPSASE\_2; UNKNOWN\_2.  
 DR PROSITE; PS00639; THIOL\_PROTEASE\_HIS; UNKNOWN\_1.  
 KW Complete proteome.  
 SQ SEQUENCE 1057 AA; 117171 MW; E3E179EF0591F0F8 CRC64;

Query Match 66.1%; Score 37; DB 16; Length 1057;  
 Best Local Similarity 60.0%; Pred. No. 77;  
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 EVVPXGMHYS 11  
 :|:| |:  
 Db 190 EIVSNGLHYS 199

RESULT 5  
 Q46486 PRELIMINARY; PRT; 208 AA.  
 AC Q46486;  
 DT 01-NOV-1996 (TREMblrel. 01, Created)  
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)  
 DT 01-MAR-2001 (TREMblrel. 16, Last annotation update)  
 DE Hypothetical 23.0 kDa protein (Gcra).  
 GN GCRA.  
 OS Corynebacterium xerosis, and  
 OS Corynebacterium striatum.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Corynebacterineae; Corynebacteriaceae;  
 OC Corynebacterium.  
 OX NCBI\_TaxID=1725, 43770;  
 RN [1]  
 RC SPECIES FROM N.A.  
 RC SPECIES=C.xerosis; STRAIN=M82B;  
 RX MEDLINE=56117603; PubMed=8559800;  
 RA Tauch A., Kassing F., Kalinowski J., Puhler A.;  
 RT "The Corynebacterium xerosis composite transposon Tn5432 consists of  
 RT two identical insertion sequences, designated IS1249, flanking the  
 RT erythromycin resistance gene ermCX.";  
 RL Plasmid 34:119-131(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=C.striatum; STRAIN=M82B;  
 RX MEDLINE=20194806; PubMed=10732668;  
 RA Tauch A., Krieff S., Kalinowski J., Puhler A.;  
 RT "The 51,409-bp R-plasmid pTP10 from the multiresistant clinical  
 RT isolate Corynebacterium striatum M82B is composed of DNA segments  
 RT initially identified in soil bacteria and in plant, animal, and human  
 RT pathogens.";  
 RL Mol. Gen. Genet. 263:1-11(2000).  
 DR EMBL; U21300; AAC95478.1; -  
 DR EMBL; AF024666; AAG03390.1; -  
 KW Hypothetical protein: Plasmid.  
 SQ SEQUENCE 208 AA; 23012 MW; F1504BE1ECD85A6 CRC64;

Query Match 64.3%; Score 36; DB 2; Length 208;  
 Best Local Similarity 50.0%; Pred. No. 22;  
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 EVVPXGMHYS 11  
 :|:| |:  
 Db 130 DVIPEGRHYA 139

RESULT 6  
 Q8YWP1 PRELIMINARY; PRT; 252 AA.  
 AC Q8YWP1;  
 DT 01-MAR-2002 (TREMblrel. 20, Created)  
 DT 01-MAR-2002 (TREMblrel. 20, Last sequence update)

DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)  
 DE Hypothetical protein Alr1563.  
 GN ALR1563.  
 OS Anabaena sp. (strain PCC 7120).  
 OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.  
 OX NCBI\_TaxID=103690;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21595285; PubMed=11759840;  
 RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,  
 RA Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,  
 RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,  
 RA Nakazaki N., Shimpo S., Sugimoto M., Takazawa M., Yamada M.,  
 RA Yasuda M., Tabata S.;  
 RT "Complete genomic sequence of the filamentous nitrogen-fixing  
 RT cyanobacterium Anabaena sp. strain PCC 7120.";  
 RL DNA Res. 8:205-213(2001).  
 DR EMBL; AP003586; BAB77929.1; -  
 KW Hypothetical protein: Complete proteome.  
 SQ SEQUENCE 252 AA; 28831 MW; 925572DA5D1CA519 CRC64;

Query Match 64.3%; Score 36; DB 16; Length 252;  
 Best Local Similarity 50.0%; Pred. No. 27;  
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMHY 10  
 :|:| |:  
 Db 235 EMIVPAGLHF 244

RESULT 7  
 Q9AVK4 PRELIMINARY; PRT; 819 AA.  
 AC Q9AVK4;  
 DT 01-JUN-2001 (TREMblrel. 17, Created)  
 DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)  
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)  
 DE SCARECROW.  
 GN PSSCR.  
 OS Pisum sativum (Garden pea).  
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Viciae; Pisum.  
 OX NCBI\_TaxID=3888;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV ALASKA;  
 RX MEDLINE=21231727; PubMed=11333309;  
 RA Sassa N., Matsushita Y., Nakamura T., Nyunoya H.;  
 RT "The Molecular Characterization and in situ Expression Pattern of Pea  
 RT SCARECROW Gene.";  
 RL Plant Cell Physiol. 42:385-394(2001).  
 DR EMBL; AB048713; BAB39155.1; -  
 DR InterPro; IPR001444; Flag\_bb\_tod.  
 DR InterPro; IPR005202; GRAS.  
 DR Pfam; PF03514; GRAS; 1.  
 DR PROSITE; PS00588; FLAGELLIN\_BB\_ROD; UNKNOWN\_1.  
 SQ SEQUENCE 819 AA; 90372 MW; 41B67BD6DC72ADFA CRC64;

Query Match 64.3%; Score 36; DB 10; Length 819;  
 Best Local Similarity 45.5%; Pred. No. 94;  
 Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMHYS 11  
 :|:| |:  
 Db 343 DDVPTSLHFS 353

RESULT 8  
 Q57489 PRELIMINARY; PRT; 139 AA.  
 AC Q57489;  
 DT 01-NOV-1996 (TREMblrel. 01, Created)

DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE DNA ligase (fragment).  
 OS Bacteroides nodosus (Dichelobacter nodosus).  
 OC Bacteria; Proteobacteria; gamma subdivision; Cardiobacteriaceae;  
 OC Dichelobacter.  
 OX NCBI\_TaxID=870;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96020672; PubMed=7476204;  
 RA Moses E.K., Good R.T., Sinistaj M., Billington S.J., Langford C.J.,  
 RA Rood J.I.;  
 RT "A multiple site-specific DNA-inversion model for the control of Omp1  
 phase and antigenic variation in Dichelobacter nodosus.";  
 RL Mol. Microbiol. 17:183-196(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96257263; PubMed=8654969;  
 RA Billington S.J., Sinistaj M., Cheetham B.F., Ayres A., Moses E.K.,  
 RA Katz M.E., Rood J.I.;  
 RT "Identification of a native Dichelobacter nodosus plasmid and  
 implications for the evolution of the vap regions.";  
 RL Gene 172:111-116(1996).  
 DR EMBL; U02462; AAB12366.1; -;  
 DR InterPro: IPR001357; BRCT.  
 DR Pfam: PF00533; BRCT; 1.  
 DR SMART: SM00292; BRCT; 1.  
 DR PROSITE: PS50172; BRCT; 1.  
 KW Ligase.  
 FT NON\_TER  
 SQ SEQUENCE 139 AA; 15052 MW; E0E110AA4B7D4708 CRC64;  
 Query Match 62.5%; Score 35; DB 2; Length 139;  
 Best Local Similarity 55.6%; Pred. No. 24;  
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
 QY 3 VVPXGMHYS 11  
 Db 21 IVPAGVHWS 29  
 RESULT 9  
 Q9PC35  
 ID Q9PC35 PRELIMINARY; PRT; 156 AA.  
 AC Q9PC35;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DE 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 DE Hypothetical protein Xf1950.  
 GN Xf1950.  
 OS Xylella fastidiosa.  
 OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;  
 OC Xylella.  
 OX NCBI\_TaxID=2371;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=9A5C;  
 RX MEDLINE=20365717; PubMed=10910347;  
 RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,  
 RA Alvaranga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,  
 RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,  
 RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carriaro D.M., Carrier H.,  
 RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,  
 RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,  
 RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,  
 RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furian L.R.,  
 RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,  
 RA Ho P.D., Hohenisei J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,  
 RA Krieger J.E., Kuranae E.E., Laigret F., Lambais M.R., Leite L.C.C.,  
 RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,  
 RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,  
 RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,  
 RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,

RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,  
 RA Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,  
 RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,  
 RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,  
 RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,  
 RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,  
 RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,  
 RA da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,  
 RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tshako M.H.,  
 RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,  
 RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;  
 RT "The genome sequence of the plant pathogen Xylella fastidiosa.";  
 RL Nature 406:151-159(2000).  
 DR EMBL; AE004014; AAF84752.1; -;  
 DR InterPro: IPR002545; Chew.  
 DR SMART: SM00260; Chew; 1.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 156 AA; 17144 MW; D8358619C6671A5D CRC64;  
 Query Match 62.5%; Score 35; DB 16; Length 156;  
 Best Local Similarity 55.6%; Pred. No. 27;  
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 EEVVPXGMH 9  
 Db 119 EEILPQGVH 127  
 RESULT 10  
 Q40479  
 ID Q40479 PRELIMINARY; PRT; 233 AA.  
 AC Q40479;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DE 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE EREBP-2.  
 OS Nicotiana tabacum (Common tobacco).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.  
 OX NCBI\_TaxID=4097;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BY4; TISSUE=LEAF;  
 RX MEDLINE=95276459; PubMed=7756828;  
 RA Olme-Takagi M., Shinshi H.;  
 RT "Ethylene-inducible DNA binding proteins that interact with an  
 ethylene responsive element.";  
 RL Plant Cell 7:173-182(1995).  
 DR EMBL; D38126; BAA07324.1; -;  
 DR HSSP; O80337; 2GCC.  
 DR TRANSFAC; T02654; -;  
 DR InterPro: IPR001471; TF\_ERF.  
 DR Pfam: PF00847; AP2-domain; 1.  
 DR PRINTS; PR00367; ETHRSPLEMT.  
 DR Prodom; PD001423; TF\_AP2; 1.  
 DR SMART; SM00380; AP2; 1.  
 SQ SEQUENCE 233 AA; 25563 MW; 6CD16783582C0CB5 CRC64;  
 Query Match 62.5%; Score 35; DB 10; Length 233;  
 Best Local Similarity 60.0%; Pred. No. 41;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 EEVVPXGMH 10  
 Db 90 QAVVPKGRHV 99  
 RESULT 11  
 Q9LW50  
 ID Q9LW50 PRELIMINARY; PRT; 237 AA.  
 AC Q9LW50;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)

DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE Ethylene-responsive element binding factor.  
 GN NSERF2.  
 OS Nicotiana sylvestris (Wood tobacco).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.  
 OX NCBI\_TaxID=4096;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20399450; PubMed=10945353;  
 RA Kitajima S., Koyama T., Ohme-Takagi M., Shinshi H., Sato F.;  
 RT "Characterization of gene expression of NSERFs, transcription factors  
 of basic PR genes from Nicotiana sylvestris.";  
 RL Plant Cell Physiol. 41:817-824(2000).  
 DR EMBL; AB016264; BAA97122.1; -.  
 DR HSSP; O80337; 2GCC.  
 DR InterPro; IPR001471; TF\_ERF.  
 DR Pfam; PF00847; AP2-domain; 1.  
 DR PRINTS; PR00367; ETHRSPLEMT.  
 DR ProDom; PD001423; TF\_AP2; 1.  
 DR SMART; SM00380; AP2; 1.  
 SQ SEQUENCE 237 AA; 26243 MW; 01BC3EEB51E46298 CRC64;  
 Query Match 62.5%; Score 35; DB 10; Length 237;  
 Best Local Similarity 60.0%; Pred. No. 41;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 EEVVPXGMHY 10  
 : ||| | |  
 Db 94 QAVVPKGRHY 103  
 RESULT 12  
 Q38317 ID Q38317 PRELIMINARY; PRT; 317 AA.  
 AC Q38317;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE Lysin.  
 GN LYS.  
 OS Lactobacillus bacteriophage phi adh.  
 OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae.  
 OX NCBI\_TaxID=12417;  
 RN [1]  
 RA Altermann E.;  
 RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99384014; PubMed=10452953;  
 RA Altermann E., Klein J., Henrich B.;  
 RT "Primary structure and features of the genome of the Lactobacillus  
 gasseri temperate bacteriophage phi-adh.";  
 RL Gene 236:333-346(1999).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95138034; PubMed=7836307;  
 RA Henrich B., Binshofer B., Blaesi U.;  
 RT "Primary structure and functional analysis of the lysis genes of  
 Lactobacillus gasseri bacteriophage phi-adh.";  
 RL J. Bacteriol. 177:723-732(1995).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93231538; PubMed=8472961;  
 RA Frenaux C., De Antoni G., Raya R., Klaenhammer T.;  
 RT "Genetic organization and sequence of the region encoding integrative  
 functions from Lactobacillus gasseri temperate bacteriophage phi-  
 adh.";  
 RL Gene 126:61-66(1993).  
 RN [5]

RP SEQUENCE FROM N.A.  
 RA Engel G., Altermann E., Klein J., Henrich B.;  
 RT "Structure of a genome region of the Lactobacillus gasseri temperate  
 phage phi adh covering a repressor gene and cognate promoters.";  
 RL Gene 210:67-70(1998).  
 DR EMBL; AJ131519; CAB52540.1; -.  
 DR InterPro; IPR002053; GH\_25.  
 DR Pfam; PF01183; Glyco\_hydro\_25; 1.  
 DR ProDom; PD004620; GH\_25; 1.  
 DR SMART; SM00287; SH3b; 1.  
 SQ SEQUENCE 317 AA; 34703 MW; 9FF2715EB43561C7 CRC64;  
 Query Match 62.5%; Score 35; DB 9; Length 317;  
 Best Local Similarity 66.7%; Pred. No. 56;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 QY 3 VVPXGMHY 11  
 : ||| | |  
 Db 60 VVPWGYHYA 68  
 RESULT 13  
 Q9XVK4 ID Q9XVK4 PRELIMINARY; PRT; 425 AA.  
 AC Q9XVK4;  
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 DE RI0D12.10 protein.  
 GN RI0D12.10.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Percy C.M.;  
 RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99069613; PubMed=9851916;  
 RA none;  
 RT "Genome sequence of the nematode C.elegans: A platform for  
 investigating biology.";  
 RL Science 282:2012-2018(1998).  
 DR EMBL; Z81109; CAB03241.1; -.  
 DR InterPro; IPR000719; Euk\_pkinase.  
 DR Pfam; PF00069; pkinase; 1.  
 DR ProDom; PD000001; Euk\_pkinase; 1.  
 DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
 KW ATP-binding; Transferase.  
 SQ SEQUENCE 425 AA; 49410 MW; 5D96E29B08C8E9D6 CRC64;  
 Query Match 62.5%; Score 35; DB 5; Length 425;  
 Best Local Similarity 50.0%; Pred. No. 76;  
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 EEVVPXGMHY 10  
 : ||| | |  
 Db 335 EQUIVPGGLQY 344  
 RESULT 14  
 Q9SA71 ID Q9SA71 PRELIMINARY; PRT; 510 AA.  
 AC Q9SA71;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE T518.1 protein.  
 GN T518.1.  
 OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC Eusids II; Brassicales; Brassicaceae; Arabidopsids.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. COLUMBIA;  
 RA Vysotskaia V.S., Schwartz J.R., Yu G., Toriumi M., Lenz C., Liu S.,  
 RA Li J., Kremenetskaia I., Luros J., Ngan I., Gonzalez A., Altafi H.,  
 RA Araujo R., Chao Q., Conn L., Conway A.B., Dunn P., Hansen N.,  
 RA Huizar L., Kim C., Palm C., Rowley D., Shinn P., Walker M.,  
 RA Davis R.W., Ecker J.R., Federspiel N.A., Theologis A.;  
 RT "Arabidopsis thaliana chromosome 1 BAC 1518 sequence.";  
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).  
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AC007060; AAD25743.1; -;  
 KW Transmembrane.  
 SQ SEQUENCE 510 AA; 55089 MW; 0DC63CC41F4712F8 CRC64;

Query Match 62.5%; Score 35; DB 10; Length 510;  
 Best Local Similarity 60.0%; Pred. No. 92;  
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 EEVVPXGMHY 10  
 Db ||||| |:  
 12 EEVKPGIHF 21

RESULT 15  
 Q87Z07 PRELIMINARY; PRT; 2042 AA.  
 AC Q87Z07;  
 DT 01-JUN-2002 (TREMBlrel. 21, Created)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)  
 DE Predicted protein of the CobN/Mg-chelatase family.  
 GN MK0134.  
 OS Methanopyrus kandleri.  
 OC Archaea; Euryarchaeota; Methanopyri; Methanopyrales; Methanopyraceae;  
 OC Methanopyrus.  
 OX NCBI\_TaxID=2320;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=AV19 / DSM 6324 / JCM 9639;  
 RX MEDLINE=21927647; PubMed=11930014;  
 RA Slesarev A.I., Mezhevaya K.V., Makarova K.S., Polushin N.N.,  
 RA Shcherbinina O.V., Shakhova V.V., Belova G.I., Aravind L.,  
 RA Natale D.A., Rogozin I.B., Tatusov R.L., Wolf Y.I., Stetter K.O.,  
 RA Malykh A.G., Koonin E.V., Kozayavkin S.A.;  
 RT "The complete genome of hyperthermophile Methanopyrus kandleri AV19  
 and monophyly of archaeal methanogens.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:4644-4649(2002).  
 DR EMBL; AE010313; AA001351.1; -;  
 KW Complete proteome.  
 SQ SEQUENCE 2042 AA; 227716 MW; 83B6E092A62C112A CRC64;

Query Match 62.5%; Score 35; DB 17; Length 2042;  
 Best Local Similarity 75.0%; Pred. No. 4e+02;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 EVVVPXGMH 9  
 Db ||||| |:  
 924 EVVPIGLH 931

RESULT 16  
 O27902 PRELIMINARY; PRT; 264 AA.  
 AC O27902;  
 DT 01-JAN-1998 (TREMBlrel. 05, Created)  
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)  
 DE Diphthine synthase.  
 GN MK0134.  
 OS Methanopyrus kandleri.  
 OC Archaea; Euryarchaeota; Methanopyri; Methanopyrales; Methanopyraceae;  
 OC Methanopyrus.  
 OX NCBI\_TaxID=2320;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=AV19 / DSM 6324 / JCM 9639;  
 RX MEDLINE=21927647; PubMed=11930014;  
 RA Slesarev A.I., Mezhevaya K.V., Makarova K.S., Polushin N.N.,  
 RA Shcherbinina O.V., Shakhova V.V., Belova G.I., Aravind L.,  
 RA Natale D.A., Rogozin I.B., Tatusov R.L., Wolf Y.I., Stetter K.O.,  
 RA Malykh A.G., Koonin E.V., Kozayavkin S.A.;  
 RT "The complete genome of hyperthermophile Methanopyrus kandleri AV19  
 and monophyly of archaeal methanogens.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:4644-4649(2002).  
 DR EMBL; AE010313; AA001351.1; -;  
 KW Complete proteome.  
 SQ SEQUENCE 2042 AA; 227716 MW; 83B6E092A62C112A CRC64;

Query Match 60.7%; Score 34; DB 16; Length 279;  
 Best Local Similarity 75.0%; Pred. No. 79;  
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 4 VPXGMHYS 11

GN MTH1874.  
 OS Methanobacterium thermoautotrophicum.  
 OC Archaea; Euryarchaeota; Methanobacteriales;  
 OC Methanobacteriaceae; Methanothermobacter.  
 OX NCBI\_TaxID=187420;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=DELTA H;  
 RX MEDLINE=98037514; PubMed=93711463;  
 RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,  
 RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,  
 RA Harrison D., Hoang L., Keagle P., Lumm W., Pothier B., Qiu D.,  
 RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,  
 RA Jiwan N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,  
 RA McDougall S., Shiner G., Goyal A., Pietrowski S., Church G.M.,  
 RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;  
 RT "Complete genome sequence of Methanobacterium thermoautotrophicum  
 strain deltaH: functional analysis and comparative genomics.";  
 RL J. Bacteriol. 179:7135-7155(1997).  
 DR EMBL; AE000940; AAB86340.1; -;  
 DR InterPro: IPR000878; Cor/por\_Mettransf.  
 DR Pfam: PF00590; TP\_methylase; 1.  
 DR TIGRFAMs; TIGR00522; dph5; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 264 AA; 28858 MW; 366BAE4E4D992C21 CRC64;

Query Match 60.7%; Score 34; DB 17; Length 264;  
 Best Local Similarity 62.5%; Pred. No. 74;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMHY 10  
 Db ||||| |:  
 235 VVPAGLHF 242

RESULT 17  
 Q9RXN9 PRELIMINARY; PRT; 279 AA.  
 AC Q9RXN9;  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DE Hypothetical protein DR0271.  
 GN DR0271.  
 OS Deinococcus radiodurans.  
 OC Bacteria; Thermus/Deinococcus group; Deinococci; Deinococcales;  
 OC Deinococcaceae; Deinococcus.  
 OX NCBI\_TaxID=1299;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=R1;  
 RX MEDLINE=20036896; PubMed=10567266;  
 RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,  
 RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,  
 RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,  
 RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zaleski C.,  
 RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,  
 RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,  
 RA Fraser C.M.;  
 RT "Genome sequence of the radioresistant bacterium Deinococcus  
 radiodurans R1.";  
 RL Science 286:1571-1577(1999).  
 DR EMBL; AE001889; AAF09867.1; -;  
 DR TIGR; DR0271; -;  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 279 AA; 31140 MW; DCEA100E0AEFE8831 CRC64;

Query Match 60.7%; Score 34; DB 16; Length 279;  
 Best Local Similarity 75.0%; Pred. No. 79;  
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 4 VPXGMHYS 11

Db 100 VPLGRHYS 107  
|| | |||

## RESULT 18

Q98H06 PRELIMINARY; PRT; 285 AA.  
AC Q98H06  
DT 01-OCT-2001 (TREMBLrel. 18, Created)  
DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE Cytochrome c1.  
GN MLL2705.  
OS Rhizobium loti (Mesorhizobium loti).  
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
OC Phyllobacteriaceae; Mesorhizobium.  
OX NCBI\_TaxID=381;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=MAFF303099;  
RX MEDLINE=21082930; PubMed=11214968;  
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,  
RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,  
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,  
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimo S., Sugimoto M.,  
RA Takeuchi C., Yamada M., Tabata S.;  
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium  
RT Mesorhizobium loti";  
RL DNA Res. 7:331-338(2000).  
DR EMBL; AP003000; BAB49770.1; -;  
DR InterPro; IPR000345; CytC\_heme\_bind.  
DR InterPro; IPR002326; Cyt\_C1.  
DR Pfam; PF02167; Cytochrome C1; 1.  
DR PRINTS; PR00603; CYTOCHROME\_C1.  
DR PROSITE; PS00190; CYTOCHROME\_C; UNKNOWN\_1.  
KW Complete proteome.  
SQ SEQUENCE 285 AA; 30961 MW; 31D9CDE2711747EE CRC64;

Query Match 60.7%; Score 34; DB 16; Length 285;

Best Local Similarity 55.68; Pred. No. 80;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 VVPXGMHYS 11  
|:| | |||  
Db 194 VPEGTHYN 202

## RESULT 19

Q9RW92 PRELIMINARY; PRT; 350 AA.  
AC Q9RW92  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE 3-dehydroquinate synthase.  
GN DR0777.  
OS Deinococcus radiodurans.  
OC Bacteria; Thermus/Deinococcus group; Deinococci; Deinococcales;  
OC Deinococcaceae; Deinococcus.  
OX NCBI\_TaxID=1299;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=R1;  
RX MEDLINE=20036896; PubMed=10567266;  
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,  
RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,  
RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,  
RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zalewski C.,  
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,  
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,  
RA Fraser C.M.;  
RT "Genome sequence of the radioresistant bacterium Deinococcus  
RT radiodurans R1.";

Query Match 60.7%; Score 34; DB 16; Length 355;

RL Science 286:1571-1577(1999).  
DR EMBL; AE001932; AAF10353.1; -;  
DR HSSP; P07547; IQS.  
DR TIGR; DR0777; -;  
DR InterPro; IPR002658; DHQ\_synthase.  
DR Pfam; PF01761; DHQ\_synthase; 1.  
KW Complete proteome.

SQ SEQUENCE 350 AA; 37121 MW; 37601D08B2FB6116 CRC64;

Query Match 60.7%; Score 34; DB 16; Length 350;  
Best Local Similarity 60.0%; Pred. No. 1e+02;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 EVVPXGMHYS 11  
|:| | |||  
Db 252 EAVAYGMHYA 261

## RESULT 20

Q9XAM3 PRELIMINARY; PRT; 355 AA.  
ID Q9XAM3  
AC Q9XAM3  
DT 01-NOV-1999 (TREMBLrel. 12, Created)  
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE Putative DNA ligase.  
GN SC06707 OR SC4C6.17C.  
OS Streptomyces coelicolor.  
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
OC Actinomycetales; Streptomyces; Streptomycetaceae; Streptomyces.  
OX NCBI\_TaxID=1902;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=A3(2);  
RA Seeger S., Harris D.;  
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=A3(2);  
RA James K.D., Parkhill J., Barrell B.G., Rajandream M.A.;  
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=A3(2);  
RX MEDLINE=97000351; PubMed=8843436;  
RA Redenbach M., Kieser H.M., Denapante D., Eichner A., Cullum J.,  
RA Kinashi H., Hopwood D.A.;  
RT "A set of ordered cosmids and a detailed genetic and physical map for  
RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";  
RL Mol. Microbiol. 21:77-96(1996).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC STRAIN=A3(2) / M145;  
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,  
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,  
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,  
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,  
RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,  
RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,  
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,  
RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,  
RA Hopwood D.A.;  
RT "Complete genome sequence of the model actinomycete Streptomyces  
RT coelicolor A3(2).";  
RL Nature 417:141-147(2002).  
DR EMBL; AL079355; CAB45581.1; -;  
DR InterPro; IPR000977; DNA\_ligase.  
DR Pfam; PF01068; DNA\_ligase; 1.  
DR PROSITE; PS00697; DNA\_LIGASE\_A1; UNKNOWN\_1.  
KW Ligase.  
SQ SEQUENCE 355 AA; 39873 MW; 7A995A55116077B1 CRC64;

Best Local Similarity 71.48; Pred. No. 1e+02; Matches 5; Conservative 1; Mismatches 0; Indels 1; Gaps 0;

QY 4 VPXGMHY 10  
:|:||||  
Db 20 IPPGMHY 26

RESULT 21  
027679

ID O27679 PRELIMINARY; PRT; 360 AA.  
AC O27679;  
DT 01-JAN-1998 (TREMBlrel. 05, Created)  
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)  
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
DE Cell division protein.  
GN MTH1642.  
OS Methanobacterium thermoautotrophicum.  
OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;  
OC Methanobacteriaceae; Methanothermobacter.  
OX NCBI\_TaxID=187420;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-DELTA H;  
RX MEDLINE=98037514; PubMed=9371463;  
RA Smith D.R., Doucette-Stamm L.A., DeLoughery C., Lee H.-M., Dubois J.,  
RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,  
RA Harrison D., Hoang L., Keagle P., Lumm W., Pothier B., Qiu D.,  
RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,  
RA Jiwani N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,  
RA McDougall S., Shimer G., Goyal A., Pietrowski S., Church G.M.,  
RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;  
RT "Complete genome sequence of Methanobacterium thermoautotrophicum  
RT deltaH: functional analysis and comparative genomics.";  
RL J. Bacteriol. 179:7135-7155(1997).  
DR EMBL; AE000923; AB86115.1;  
DR InterPro; IPR005140; eRF1.1;  
DR InterPro; IPR005141; eRF1.2;  
DR InterPro; IPR005142; eRF1.3;  
DR InterPro; IPR004405; PELA.  
DR Pfam; PF03463; eRF1.1;  
DR Pfam; PF03464; eRF1.2;  
DR Pfam; PF03465; eRF1.3;  
DR TIGRFAMS; TIGR00111; pELA; 1.  
KW Cell division; Complete proteome.

SQ SEQUENCE 360 AA; 40814 MW; 2A000CB4B3CEF469 CRC64;

Query Match 60.7%; Score 34; DB 17; Length 360;  
Best Local Similarity 45.5%; Pred. No. 1e+02; Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMHYS 11  
|:|:|:|:|:  
Db 98 EDLVPMSGSHHT 108

RESULT 22  
Q9KBAL

ID Q9KBAL PRELIMINARY; PRT; 425 AA.  
AC Q9KBAL;  
DT 01-OCT-2000 (TREMBlrel. 15, Created)  
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
DE Hypothetical protein BH2027.  
GN BH2027.  
OS Bacillus halodurans.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;  
OC Bacillaceae; Bacillus.  
OX NCBI\_TaxID=86665;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-C-125 / JCM 9153;  
RX MEDLINE=20512582; PubMed=11058132;

RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,  
RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,  
RA Horikoshi K.;  
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus  
RT halodurans and genomic sequence comparison with Bacillus subtilis.";  
RL Nucleic Acids Res. 28:4317-4331(2000).  
DR EMBL; AP001514; BAB05746.1;  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 425 AA; 47639 MW; 914824B1FB65E3DE CRC64;

Query Match 60.7%; Score 34; DB 16; Length 425;  
Best Local Similarity 60.0%; Pred. No. 1.2e+02; Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMHY 10  
|:|:|:|:|:  
Db 403 EELLIEGMHY 412

RESULT 23

ID O16912 PRELIMINARY; PRT; 495 AA.  
AC O16912;  
DT 01-JAN-1998 (TREMBlrel. 05, Created)  
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)  
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
DE F10D2.3 protein.  
GN F10D2.3.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Pterodermidae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BRISTOL N2;  
RX MEDLINE=94150718; PubMed=7906398;  
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,  
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,  
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,  
RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,  
RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,  
RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,  
RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Shownkeen R.,  
RA Smaildon N., Smith A., Sonhammer E., Staden R., Sulston J.,  
RA Thierry-Mieg J., Thomas K., Vaubin M., Vaughan K., Waterston R.,  
RA Watery A., Weinstock L., Wilkinson-Sproat J., Wohldman P.;  
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
RT elegans.";  
RL Nature 368:32-38(1994).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BRISTOL N2;  
RA Graves T., Wohldmann P., Gillam B.;  
RT "The sequence of C. elegans cosmid F10D2.";  
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF022972; AAC48234.1;  
DR InterPro; IPR004151; Sre.  
DR Pfam; PF03125; Sre; 1.  
SQ SEQUENCE 495 AA; 58190 MW; 0C61139C138EEB4C CRC64;

Query Match 60.7%; Score 34; DB 5; Length 495;  
Best Local Similarity 50.0%; Pred. No. 1.4e+02; Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMHY 10  
|:|:|:|:|:  
Db 218 ENIVPTGKHH 227

RESULT 24

Q9Y0Y6 PRELIMINARY; PRT; 637 AA.  
ID Q9Y0Y6  
AC Q9Y0Y6;

DT 01-NOV-1999 (Tremblrel. 12, Created)  
 DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)  
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)  
 DE BCDA:LD28657 Protein.  
 DE BCDA:LD28657 OR CGI098.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BERKELEY;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
 Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos B.L.G.,  
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jialali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector R., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster.";  
 RL Science 287:2185-2195(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Rubin G.M., Wan K.H., Harvey D., Lewis S.E., Brokstein P., Tsang G.,  
 RA Agbayani A., Arcaina T., Baxter E., Blazej R.G., Butenhoff C.,  
 RA Champe M., Chavez C., Chew M., Doyle C.M., Farfan D.E., Frise E.,  
 RA Galle R., George R.A., Harris N.L., Hoskins R.A., Evans-Holm M.,  
 RA Houston K.A., Hummasti S., Kim E., Li P., Moshrefi M., Pacleb J.M.,  
 RA Park S., Sequeira A., Sethi H., Snir E., Svirskas R.R., Weinburg T.,  
 RA Celnik S.E.;  
 RT "Full length Drosophila melanogaster cDNA sequence.";  
 RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AE003601; AAF51961.1; -;  
 DR EMBL: AF145690; AAD38665.1; -;  
 DR FlyBase: FBgn0027497; BcDNA:LD28657.  
 DR InterPro: IPR000719; Euk\_pkinase.  
 DR Pfam: PF00069; pkinase; 1  
 DR ProDom: PD000001; Euk\_pkinase; 1.  
 DR PROSITE: PS50011; PROTEIN\_KINASE\_DOM; 1.  
 KW ATP-binding; Transferase.  
 SQ SEQUENCE 637 AA; 70507 MW; ABB262CAFA44D20 CRC64;

Query Match 60.7%; Score 34; DB 5; Length 637;  
 Best Local Similarity 66.7%; Pred. No. 1.9e+02;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 VVPXGMHYS 11  
 ||| :|||  
 Db 263 VVPDAVHYS 271

## RESULT 25

Q9E1X6 PRELIMINARY; PRT; 678 AA.  
 AC Q9E1X6;  
 DT 01-MAR-2001 (Tremblrel. 16, Created)  
 DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)  
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)  
 DE Hypothetical 75.9 kDa protein.  
 OS Cercopithecine herpesvirus 7.  
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
 OC Alphaherpesvirinae; Varicellovirus.  
 OX NCBI\_TaxID=35245;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Gray W.L., Starnes H.B., White M.W., Ashburn C.V., Mahalingam R.;  
 RT "Complete Sequence of the Simian Varicella Virus Genome.";  
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF275348; AAG27217.1; -;  
 KW Hypothetical protein.  
 SQ SEQUENCE 678 AA; 75850 MW; A17B09E30512FE3C CRC64;

Query Match 60.7%; Score 34; DB 12; Length 678;  
 Best Local Similarity 50.0%; Pred. No. 2e+02;  
 Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMHY 10  
 ||::|||  
 Db 147 EEIIPKGTGY 156

Search completed: June 10, 2003, 13:46:29  
 Job time : 28.7857 secs





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model  
Run on: June 10, 2003, 13:24:45 ; Search time 31.3571 Seconds  
(without alignments)  
46.744 Million cell updates/sec

Title: US-09-909-164-8  
Perfect score: 54  
Sequence: 1 BEVVPXGMDYS 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 75 summaries

Database :			
A_Geneseq_101002:*			
1:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*		
2:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*		
3:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*		
4:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*		
5:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*		
6:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*		
7:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*		
8:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*		
9:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*		
10:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*		
11:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*		
12:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*		
13:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*		
14:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*		
15:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*		
16:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*		
17:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*		
18:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*		
19:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*		
20:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*		
21:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*		
22:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*		
23:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*		

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Query Match	Length	ID	Description
1	52	96.3	11	23	ABB80524	Hepatitis C virus
2	52	96.3	11	23	ABB80528	Hepatitis C virus
3	52	96.3	11	23	ABB80529	Hepatitis C virus
4	52	96.3	11	23	ABB80561	Hepatitis C virus
5	52	96.3	11	23	ABB80562	Hepatitis C virus
6	47	87.0	11	23	ABB80538	Hepatitis C virus
7	47	87.0	11	23	ABB80542	Hepatitis C virus
8	47	87.0	11	23	ABB80543	Hepatitis C virus
9	46	85.2	11	23	ABB80521	Hepatitis C virus
10	46	85.2	11	23	ABB80522	Hepatitis C virus

11	46	85.2	11	23	ABB80525	Hepatitis C virus
12	46	85.2	11	23	ABB80526	Hepatitis C virus
13	46	85.2	11	23	ABB80547	Hepatitis C virus
14	46	85.2	11	23	ABB80548	Hepatitis C virus
15	46	85.2	11	23	ABB80551	Hepatitis C virus
16	46	85.2	11	23	ABB80556	Hepatitis C virus
17	46	85.2	11	23	ABB80557	Hepatitis C virus
18	46	85.2	11	23	ABB80559	Hepatitis C virus
19	46	85.2	11	23	ABB80563	Hepatitis C virus
20	46	85.2	11	23	ABB80564	Hepatitis C virus
21	46	85.2	11	23	ABB80565	Hepatitis C virus
22	46	85.2	11	23	ABB80566	Hepatitis C virus
23	46	85.2	11	23	ABB80567	Hepatitis C virus
24	46	85.2	11	23	ABB80568	Hepatitis C virus
25	45	83.3	11	23	ABB80523	Hepatitis C virus
26	45	83.3	11	23	ABB80527	Hepatitis C virus
27	45	83.3	11	23	ABB80538	Hepatitis C virus
28	45	83.3	11	23	ABB80560	Hepatitis C virus
29	44	81.5	11	23	ABB80533	Hepatitis C virus
30	44	81.5	11	23	ABB80534	Hepatitis C virus
31	41	75.9	11	23	ABB80535	Hepatitis C virus
32	41	75.9	11	23	ABB80536	Hepatitis C virus
33	41	75.9	11	23	ABB80539	Hepatitis C virus
34	41	75.9	11	23	ABB80540	Hepatitis C virus
35	40	74.1	11	23	ABB80537	Hepatitis C virus
36	40	74.1	11	23	ABB80541	Hepatitis C virus
37	40	74.1	11	23	ABB80544	Hepatitis C virus
38	40	74.1	11	23	ABB80545	Hepatitis C virus
39	40	74.1	11	23	ABB80549	Hepatitis C virus
40	40	74.1	11	23	ABB80552	Hepatitis C virus
41	40	74.1	11	23	ABB80553	Hepatitis C virus
42	39	72.2	11	23	ABB80546	Hepatitis C virus
43	39	72.2	11	23	ABB80550	Hepatitis C virus
44	39	72.2	11	23	ABB80554	Hepatitis C virus
45	39	72.2	11	23	ABB80555	Hepatitis C virus
46	37	68.5	11	23	ABB80530	Hepatitis C virus
47	37	68.5	11	23	ABB80531	Hepatitis C virus
48	37	68.5	11	23	ABB80532	Hepatitis C virus
49	36	66.7	20	20	AAU76810	Hepatitis C virus
50	36	66.7	34	22	AAE00907	Citrus partial suc
51	36	66.7	34	22	AAE00908	Citrus partial suc
52	36	66.7	460	23	ABG91579	Herbicidally activ
53	36	66.7	1022	22	ABG03621	Novel human diagno
54	36	66.7	1022	22	ABG05826	Novel human diagno
55	36	66.7	1022	22	ABG08173	Novel human diagno
56	36	66.7	1068	13	AAK20198	Sucrose phosphate
57	36	66.7	1068	13	AAK27931	SPS protein. Zea
58	36	66.7	1068	18	AAW38266	Maize sucrose phos
59	36	66.7	1068	18	AAW09869	Sucrose phosphate
60	36	66.7	1071	15	AAK60796	Rice sucrose phosph
61	36	66.7	1083	23	ABG92875	Herbicidally activ
62	35	64.8	2778	22	ABG58683	Drosophila melanog
63	34.5	63.9	748	22	AAW40225	Human polypeptide
64	34	63.0	140	16	AAK64232	MAB L243 VH region
65	34	63.0	140	16	AAK64235	Humanized antibody
66	34	63.0	140	16	AAK64265	CDR-grafted L243-g
67	34	63.0	140	16	AAK64257	MHC-II MAB L243 he
68	34	63.0	143	22	ABG60256	Drosophila melanog
69	34	63.0	150	21	ABG28379	Arabidopsis thalia
70	34	63.0	150	22	AAU87623	Bovine mammary tis
71	34	63.0	222	22	AAU03629	Group B Streptococ
72	34	63.0	222	23	ABP26468	Streptococcus poly
73	34	63.0	565	23	ABG53631	Lactococcus lactis
74	34	63.0	587	21	AAV74287	Neisseria meningit
75	34	63.0	1062	22	AAU33615	Pseudomonas aerugi

ALIGNMENTS

RESULT 1  
ABB80524  
ID ABB80524 standard; peptide; 11 AA.

XX ABB80524;  
 AC  
 DT 08-OCT-2002 (first entry)  
 XX  
 DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #4.  
 XX  
 KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
 KW virucide.  
 XX  
 OS Synthetic.  
 XX  
 PH Key Location/Qualifiers  
 FT Modified-site 1 /note= "N-terminal acetyl"  
 FT Modified-site 6  
 FT /note= "Norvalyl carbonyl forming keto-amide linkage with  
 FT residue 7"  
 FT  
 FT Misc-difference 8  
 FT /note= "D-form residue"  
 FT Modified-site 11  
 FT /note= "C-terminal amide"  
 FT  
 FT WO200208251-A2.  
 FT  
 FT 31-JAN-2002.  
 FT  
 FT 19-JUL-2001; 2001WO-US23169.  
 FT  
 FT 21-JUL-2000; 2000US-220101P.  
 FT  
 FT (CORV-) CORVAS INT INC.  
 FT  
 FT Lim-wilby M, Levy OE, Brunck TK;  
 FT  
 FT WPI; 2002-361643/39.  
 FT  
 FT Novel peptide compound having hepatitis C virus protease inhibitory  
 FT activity useful for treating disorders associated with hepatitis C  
 FT virus protease  
 FT  
 FT Claim 17; Page 64; 69pp; English.  
 FT  
 FT The sequence represents a peptide compound of the invention having  
 FT hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
 FT invention are alpha-ketoamide peptide analogues. The peptides have  
 FT virucide activity, and are useful for treating and in the manufacture of  
 FT a medicament to treat disorders associated with HCV protease. A  
 FT pharmaceutical composition comprising the peptide as an active ingredient  
 FT is useful for treating disorders associated with hepatitis C virus.  
 FT  
 FT Sequence 11 AA;  
 FT  
 FT Query Match 96.3%; Score 52; DB 23; Length 11;  
 FT Best Local Similarity 100.0%; Pred. No. 0.0015;  
 FT Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 FT  
 FT QY 1 EEVVPXGMDYS 11  
 FT | | | | | | | | | |  
 FT Db 1 EEVVPXGMDYS 11  
 FT  
 FT RESULT 2  
 FT ABB80528  
 FT ID ABB80528 standard; peptide; 11 AA.  
 FT  
 FT AC ABB80528;  
 FT  
 FT 08-OCT-2002 (first entry)  
 FT  
 FT Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #8.  
 FT  
 FT Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
 FT  
 FT KW

KW virucide.  
 XX  
 OS Synthetic.  
 XX  
 PH Key Location/Qualifiers  
 FT Modified-site 1 /note= "N-terminal acetyl"  
 FT Modified-site 6  
 FT /note= "Norvalyl carbonyl forming keto-amide linkage with  
 FT residue 7"  
 FT  
 FT Misc-difference 8  
 FT /note= "D-form residue"  
 FT Modified-site 11  
 FT /note= "C-terminal amide"  
 FT  
 FT WO200208251-A2.  
 FT  
 FT 31-JAN-2002.  
 FT  
 FT 19-JUL-2001; 2001WO-US23169.  
 FT  
 FT 21-JUL-2000; 2000US-220101P.  
 FT  
 FT (CORV-) CORVAS INT INC.  
 FT  
 FT Lim-wilby M, Levy OE, Brunck TK;  
 FT  
 FT WPI; 2002-361643/39.  
 FT  
 FT Novel peptide compound having hepatitis C virus protease inhibitory  
 FT activity useful for treating disorders associated with hepatitis C  
 FT virus protease  
 FT  
 FT Claim 17; Page 64; 69pp; English.  
 FT  
 FT The sequence represents a peptide compound of the invention having  
 FT hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
 FT invention are alpha-ketoamide peptide analogues. The peptides have  
 FT virucide activity, and are useful for treating and in the manufacture of  
 FT a medicament to treat disorders associated with HCV protease. A  
 FT pharmaceutical composition comprising the peptide as an active ingredient  
 FT is useful for treating disorders associated with hepatitis C virus.  
 FT  
 FT Sequence 11 AA;  
 FT  
 FT Query Match 96.3%; Score 52; DB 23; Length 11;  
 FT Best Local Similarity 100.0%; Pred. No. 0.0015;  
 FT Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 FT  
 FT QY 1 EEVVPXGMDYS 11  
 FT | | | | | | | | | |  
 FT Db 1 EEVVPXGMDYS 11  
 FT  
 FT RESULT 3  
 FT ABB80529  
 FT ID ABB80529 standard; peptide; 11 AA.  
 FT  
 FT AC ABB80529;  
 FT  
 FT 08-OCT-2002 (first entry)  
 FT  
 FT Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #9.  
 FT  
 FT Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
 FT virucide.  
 FT  
 FT OS Synthetic.  
 FT  
 FT PH Key Location/Qualifiers  
 FT Modified-site 1 /note= "N-terminal acetyl"  
 FT Modified-site 6  
 FT

FT FT /note= "Norvalyl carbonyl forming keto-amide linkage with  
FT FT residue 7"  
FT FT  
FT FT Misc-difference 8 /note= "D-form residue"  
FT FT Misc-difference 9 /note= "D-form residue"  
FT FT Modified-site 11 /note= "C-terminal amide"  
FT FT  
XX XX WO200208251-A2.  
XX XX 31-JAN-2002.  
XX XX 19-JUL-2001; 2001WO-US23169.  
XX XX 21-JUL-2000; 2000US-220101P.  
XX XX (CORV-) CORVAS INT INC.  
XX XX Lim-wilby M, Levy OE, Brunck TK;  
XX XX WPI; 2002-361643/39.  
XX XX Novel peptide compound having hepatitis C virus protease inhibitory  
XX XX activity useful for treating disorders associated with hepatitis C  
XX XX virus protease  
XX XX Claim 17; Page 64; 69pp; English.  
XX XX The sequence represents a peptide compound of the invention having  
XX XX hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
XX XX invention are alpha-ketoamide peptide analogues. The peptides have  
XX XX virucide activity, and are useful for treating and in the manufacture of  
XX XX a medicament to treat disorders associated with HCV protease. A  
XX XX pharmaceutical composition comprising the peptide as an active ingredient  
XX XX is useful for treating disorders associated with hepatitis C virus.  
SQ Sequence 11 AA;  
Query Match 96.3%; Score 52; DB 23; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.0015;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 EEVVPXGMDYS 11  
DB 1 EEVVPXGMDYS 11  
RESULT 4  
ABB80561  
ID ABB80561 standard; peptide; 11 AA.  
XX AC ABB80561;  
XX XX  
XX XX 08-OCT-2002 (first entry)  
XX XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #41.  
XX XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
XX XX virucide.  
XX XX Synthetic.  
XX XX Key Location/Qualifiers  
FT FT Modified-site 1 /note= "N-terminal acetyl"  
FT FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with  
FT FT residue 7"  
FT FT Misc-difference 8 /note= "D-form residue"  
FT FT Modified-site 8 /note= "Oxymethionine"  
FT FT

FT FT Modified-site 11 /note= "C-terminal amide"  
FT FT  
XX XX WO200208251-A2.  
XX XX 31-JAN-2002.  
XX XX 19-JUL-2001; 2001WO-US23169.  
XX XX 21-JUL-2000; 2000US-220101P.  
XX XX (CORV-) CORVAS INT INC.  
XX XX Lim-wilby M, Levy OE, Brunck TK;  
XX XX WPI; 2002-361643/39.  
XX XX Novel peptide compound having hepatitis C virus protease inhibitory  
XX XX activity useful for treating disorders associated with hepatitis C  
XX XX virus protease  
XX XX Claim 17; Page 65; 69pp; English.  
XX XX The sequence represents a peptide compound of the invention having  
XX XX hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
XX XX invention are alpha-ketoamide peptide analogues. The peptides have  
XX XX virucide activity, and are useful for treating and in the manufacture of  
XX XX a medicament to treat disorders associated with HCV protease. A  
XX XX pharmaceutical composition comprising the peptide as an active ingredient  
XX XX is useful for treating disorders associated with hepatitis C virus.  
SQ Sequence 11 AA;  
Query Match 96.3%; Score 52; DB 23; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.0015;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 EEVVPXGMDYS 11  
DB 1 EEVVPXGMDYS 11  
RESULT 5  
ABB80562  
ID ABB80562 standard; peptide; 11 AA.  
XX AC ABB80562;  
XX XX  
XX XX 08-OCT-2002 (first entry)  
XX XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #42.  
XX XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
XX XX virucide.  
XX XX Synthetic.  
XX XX Key Location/Qualifiers  
FT FT Modified-site 1 /note= "N-terminal acetyl"  
FT FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with  
FT FT residue 7"  
FT FT Misc-difference 8 /note= "D-form residue"  
FT FT Modified-site 8 /note= "Oxymethionine"  
FT FT Misc-difference 9 /note= "D-form residue"  
FT FT Modified-site 11 /note= "C-terminal amide"  
XX XX WO200208251-A2.  
PN

```

XX PD 31-JAN-2002.
XX PF
XX PR 19-JUL-2001; 2001WO-US23169.
XX PR 21-JUL-2000; 2000US-220101P.
XX PA (CORV-) CORVAS INT INC.
XX PI Lim-wilby M, Levy OE, Brunck TK;
XX DR WPI; 2002-361643/39.
XX PT Novel peptide compound having hepatitis C virus protease inhibitory
XX PT activity useful for treating disorders associated with hepatitis C
XX PT virus protease
XX PS Claim 17; Page 65; 69pp; English.
XX CC The sequence represents a peptide compound of the invention having
XX CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
XX CC invention are alpha-ketoamide peptide analogues. The peptides have
XX CC virucide activity, and are useful for treating and in the manufacture of
XX CC a medicament to treat disorders associated with HCV protease. A
XX CC pharmaceutical composition comprising the peptide as an active ingredient
XX CC is useful for treating disorders associated with hepatitis C virus.
XX SQ Sequence 11 AA;
XX Query Match 96.3%; Score 52; DB 23; Length 11;
XX Best Local Similarity 100.0%; Pred. No. 0.0015;
XX Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX QY 1 EEVVPXGMDYS 11
XX DB |||||
XX 1 EEVVPXGMDYS 11
XX RESULT 6
XX ABB80538
XX ID ABB80538 standard; peptide; 11 AA.
XX AC ABB80538;
XX XX
XX DT 08-OCT-2002 (first entry)
XX DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #18.
XX KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
XX KW virucide.
XX OS Synthetic.
XX XX
XX FH Key Location/Qualifiers
XX FT Modified-site 1 /note= "N-terminal acetyl"
XX FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
XX FT residue 7"
XX FT Misc-difference 9
XX FT Modified-site 11 /note= "D-form residue"
XX FT Modified-site 11 /note= "C-terminal amide"
XX PN WO200208251-A2.
XX PD 31-JAN-2002.
XX PF 19-JUL-2001; 2001WO-US23169.
XX PR 21-JUL-2000; 2000US-220101P.
XX PA (CORV-) CORVAS INT INC.

```

```

XX PI Lim-wilby M, Levy OE, Brunck TK;
XX DR WPI; 2002-361643/39.
XX XX
XX PT Novel peptide compound having hepatitis C virus protease inhibitory
XX PT activity useful for treating disorders associated with hepatitis C
XX PT virus protease
XX PS Claim 17; Page 64; 69pp; English.
XX XX
XX CC The sequence represents a peptide compound of the invention having
XX CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
XX CC invention are alpha-ketoamide peptide analogues. The peptides have
XX CC virucide activity, and are useful for treating and in the manufacture of
XX CC a medicament to treat disorders associated with HCV protease. A
XX CC pharmaceutical composition comprising the peptide as an active ingredient
XX CC is useful for treating disorders associated with hepatitis C virus.
XX SQ Sequence 11 AA;
XX Query Match 87.0%; Score 47; DB 23; Length 11;
XX Best Local Similarity 90.9%; Pred. No. 0.014;
XX Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX QY 1 EEVVPXGMDYS 11
XX DB |||||
XX 1 EEVVPXGMDYS 11
XX RESULT 7
XX ABB80542
XX ID ABB80542 standard; peptide; 11 AA.
XX AC ABB80542;
XX XX
XX DT 08-OCT-2002 (first entry)
XX DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #22.
XX KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
XX KW virucide.
XX OS Synthetic.
XX XX
XX FH Key Location/Qualifiers
XX FT Modified-site 1 /note= "N-terminal acetyl"
XX FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
XX FT residue 7"
XX FT Misc-difference 8
XX FT Modified-site 11 /note= "D-form residue"
XX FT Modified-site 11 /note= "C-terminal amide"
XX PN WO200208251-A2.
XX PD 31-JAN-2002.
XX PF 19-JUL-2001; 2001WO-US23169.
XX PR 21-JUL-2000; 2000US-220101P.
XX PA (CORV-) CORVAS INT INC.
XX PI Lim-wilby M, Levy OE, Brunck TK;
XX DR WPI; 2002-361643/39.
XX XX
XX PT Novel peptide compound having hepatitis C virus protease inhibitory
XX PT activity useful for treating disorders associated with hepatitis C
XX PT virus protease

```

XX Claim 17; Page 65; 69pp; English.  
 PS The sequence represents a peptide compound of the invention having  
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
 CC invention are alpha-ketoamide peptide analogues. The peptides have  
 CC virucide activity, and are useful for treating and in the manufacture of  
 CC a medicament to treat disorders associated with HCV protease. A  
 CC pharmaceutical composition comprising the peptide as an active ingredient  
 CC is useful for treating disorders associated with hepatitis C virus.  
 XX SQ Sequence 11 AA;  
 Query Match 87.0%; Score 47; DB 23; Length 11;  
 Best Local Similarity 90.9%; Pred. No. 0.014; Indels 0; Gaps 0;  
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 EEVVPXGMDYS 11  
 Db 1 EEVVPXGQDYS 11  
 RESULT 8  
 ABB80543  
 ID ABB80543 standard; peptide; 11 AA.  
 AC ABB80543;  
 XX 08-OCT-2002 (first entry)  
 DT Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #23.  
 DE Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
 KW virucide.  
 KW Synthetic.  
 OS Key  
 XX Location/Qualifiers  
 FH Modified-site 1 /note= "N-terminal acetyl"  
 FT Modified-site 6  
 FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with  
 FT residue 7"  
 FT Misc-difference 8 /note= "D-form residue"  
 FT Misc-difference 9 /note= "D-form residue"  
 FT Modified-site 11 /note= "C-terminal amide"  
 FT WO200208251-A2.  
 PN 31-JAN-2002.  
 XX 19-JUL-2001; 2001WO-US23169.  
 XX 21-JUL-2000; 2000US-220101P.  
 XX (CORV-) CORVAS INT INC.  
 XX Lim-wilby M, Levy OE, Brunck TK;  
 XX WPI; 2002-361643/39.  
 XX Novel peptide compound having hepatitis C virus protease inhibitory  
 PT activity useful for treating disorders associated with hepatitis C  
 PT virus protease  
 XX Claim 17; Page 65; 69pp; English.  
 PS The sequence represents a peptide compound of the invention having  
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
 CC invention are alpha-ketoamide peptide analogues. The peptides have

CC virucide activity, and are useful for treating and in the manufacture of  
 CC a medicament to treat disorders associated with HCV protease. A  
 CC pharmaceutical composition comprising the peptide as an active ingredient  
 CC is useful for treating disorders associated with hepatitis C virus.  
 XX SQ Sequence 11 AA;  
 Query Match 87.0%; Score 47; DB 23; Length 11;  
 Best Local Similarity 90.9%; Pred. No. 0.014; Indels 0; Gaps 0;  
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 EEVVPXGMDYS 11  
 Db 1 EEVVPXGQDYS 11  
 RESULT 9  
 ABB80521  
 ID ABB80521 standard; peptide; 11 AA.  
 AC ABB80521;  
 XX 08-OCT-2002 (first entry)  
 DT Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #1.  
 DE Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
 KW virucide.  
 KW Synthetic.  
 OS Key  
 XX Location/Qualifiers  
 FH Modified-site 1 /note= "N-terminal acetyl"  
 FT Modified-site 6  
 FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with  
 FT residue 7"  
 FT Modified-site 11 /note= "C-terminal amide"  
 FT WO200208251-A2.  
 PN 31-JAN-2002.  
 XX 19-JUL-2001; 2001WO-US23169.  
 XX 21-JUL-2000; 2000US-220101P.  
 XX (CORV-) CORVAS INT INC.  
 XX Lim-wilby M, Levy OE, Brunck TK;  
 XX WPI; 2002-361643/39.  
 XX Novel peptide compound having hepatitis C virus protease inhibitory  
 PT activity useful for treating disorders associated with hepatitis C  
 PT virus protease  
 XX Claim 17; Page 64; 69pp; English.  
 PS The sequence represents a peptide compound of the invention having  
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
 CC invention are alpha-ketoamide peptide analogues. The peptides have  
 CC virucide activity, and are useful for treating and in the manufacture of  
 CC a medicament to treat disorders associated with HCV protease. A  
 CC pharmaceutical composition comprising the peptide as an active ingredient  
 CC is useful for treating disorders associated with hepatitis C virus.  
 XX SQ Sequence 11 AA;  
 Query Match 85.2%; Score 46; DB 23; Length 11;  
 Best Local Similarity 90.9%; Pred. No. 0.022; Indels 1; Gaps 0;  
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXGMDYS 11  
 |||||  
 Db 1 EEVVPXGMSYS 11

## RESULT 10

ABB80522

ID ABB80522 standard; peptide; 11 AA.

XX AC ABB80522;

XX DT 08-OCT-2002 (first entry)

XX DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #2.

XX KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;

XX KW virucide.

XX OS Synthetic.

XX FH Key Location/Qualifiers

FT Modified-site 1 /note= "N-terminal acetyl"

FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with

FT residue 7"

FT Misc-difference 9 /note= "D-form residue"

FT Modified-site 11 /note= "C-terminal amide"

FT

XX WO200208251-A2.

XX 31-JAN-2002.

XX 19-JUL-2001; 2001WO-US23169.

XX 21-JUL-2000; 2000US-220101P.

XX (CORV-) CORVAS INT INC.

XX Lim-wilby M, Levy OE, Brunck TK;

XX WPI; 2002-361643/39.

XX Novel peptide compound having hepatitis C virus protease inhibitory

PT activity useful for treating disorders associated with hepatitis C

PT virus protease

XX Claim 17; Page 64; 69pp; English.

XX The sequence represents a peptide compound of the invention having

CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the

CC invention are alpha-ketoamide peptide analogues. The peptides have

CC virucide activity, and are useful for treating and in the manufacture of

CC a medicament to treat disorders associated with HCV protease. A

CC pharmaceutical composition comprising the peptide as an active ingredient

CC is useful for treating disorders associated with hepatitis C virus.

XX SQ Sequence 11 AA;

Query Match 85.2%; Score 46; DB 23; Length 11;

Best Local Similarity 90.9%; Pred. No. 0.022;

Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXGMDYS 11

|||||

Db 1 EEVVPXGMSYS 11

RESULT 11

ABB80522

XX AC ABB80522;

XX DT 08-OCT-2002 (first entry)

XX DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #6.

XX

ID ABB80525 standard; peptide; 11 AA.

XX AC ABB80525;

XX DT 08-OCT-2002 (first entry)

XX DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #5.

XX KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;

XX KW virucide.

XX OS Synthetic.

XX FH Key Location/Qualifiers

FT Modified-site 1 /note= "N-terminal acetyl"

FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with

FT residue 7"

FT Misc-difference 8 /note= "D-form residue"

FT Modified-site 11 /note= "C-terminal amide"

FT

XX WO200208251-A2.

XX 31-JAN-2002.

XX 19-JUL-2001; 2001WO-US23169.

XX 21-JUL-2000; 2000US-220101P.

XX (CORV-) CORVAS INT INC.

XX Lim-wilby M, Levy OE, Brunck TK;

XX WPI; 2002-361643/39.

XX Novel peptide compound having hepatitis C virus protease inhibitory

PT activity useful for treating disorders associated with hepatitis C

PT virus protease

XX Claim 17; Page 64; 69pp; English.

XX The sequence represents a peptide compound of the invention having

CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the

CC invention are alpha-ketoamide peptide analogues. The peptides have

CC virucide activity, and are useful for treating and in the manufacture of

CC a medicament to treat disorders associated with HCV protease. A

CC pharmaceutical composition comprising the peptide as an active ingredient

CC is useful for treating disorders associated with hepatitis C virus.

XX SQ Sequence 11 AA;

Query Match 85.2%; Score 46; DB 23; Length 11;

Best Local Similarity 90.9%; Pred. No. 0.022;

Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXGMDYS 11

|||||

Db 1 EEVVPXGMSYS 11

RESULT 12

ABB80526

ID ABB80526 standard; peptide; 11 AA.

XX AC ABB80526;

XX DT 08-OCT-2002 (first entry)

XX DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #6.

XX

KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
 KW virucide.  
 XX  
 OS Synthetic.  
 XX  
 FH Location/Qualifiers  
 FT Modified-site 1 /note= "N-terminal acetyl"  
 FT Modified-site 6  
 FT Modified-site 11 /note= "Norvalyl carbonyl forming keto-amide linkage with  
 residue 7"  
 FT  
 FT Misc-difference 8  
 FT Misc-difference 9 /note= "D-form residue"  
 FT Misc-difference 11 /note= "D-form residue"  
 FT Modified-site 11 /note= "C-terminal amide"  
 FT  
 XX WO200208251-A2.  
 XX  
 PN 31-JAN-2002.  
 PD  
 XX 19-JUL-2001; 2001WO-US23169.  
 XX  
 XX 21-JUL-2000; 2000US-220101P.  
 XX  
 XX (CORV-) CORVAS INT INC.  
 XX  
 XX Lim-wilby M, Levy OE, Brunck TK;  
 XX WPI; 2002-361643/39.  
 XX  
 XX Novel peptide compound having hepatitis C virus protease inhibitory  
 PT activity useful for treating disorders associated with hepatitis C  
 PT virus protease  
 XX  
 XX Claim 17; Page 64; 69pp; English.  
 XX  
 XX The sequence represents a peptide compound of the invention having  
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
 CC invention are alpha-ketoamide peptide analogues. The peptides have  
 CC virucide activity, and are useful for treating and in the manufacture of  
 CC a medicament to treat disorders associated with HCV protease. A  
 CC pharmaceutical composition comprising the peptide as an active ingredient  
 CC is useful for treating disorders associated with hepatitis C virus.  
 XX  
 XX Sequence 11 AA;

Query Match 85.2%; Score 46; DB 23; Length 11;  
 Best Local Similarity 90.9%; Pred. No. 0.022;  
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 EEVVPXGMDYS 11  
 |||||  
 DB 1 EEVVPXGMSYS 11

RESULT 13  
 ABB80547  
 ID ABB80547 standard; peptide; 11 AA.  
 XX  
 XX ABB80547;  
 XX  
 XX 08-OCT-2002 (first entry)  
 DT  
 XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #27.  
 DE  
 XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
 KW virucide.  
 KW  
 XX Synthetic.  
 OS  
 XX Key Location/Qualifiers  
 FH Modified-site 1 /note= "N-terminal acetyl"  
 FT Modified-site 6  
 FT Modified-site 11 /note= "Norvalyl carbonyl forming keto-amide linkage with  
 residue 7"  
 FT Misc-difference 9 /note= "D-form residue"  
 FT Modified-site 11 /note= "C-terminal amide"  
 FT  
 XX

FT Modified-site 1 /note= "N-terminal acetyl"  
 FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with  
 residue 7"  
 FT Modified-site 11 /note= "C-terminal amide"  
 FT  
 XX WO200208251-A2.  
 XX  
 PN 31-JAN-2002.  
 PD  
 XX 19-JUL-2001; 2001WO-US23169.  
 XX  
 XX 21-JUL-2000; 2000US-220101P.  
 XX  
 XX (CORV-) CORVAS INT INC.  
 XX  
 XX Lim-wilby M, Levy OE, Brunck TK;  
 XX WPI; 2002-361643/39.  
 XX  
 XX Novel peptide compound having hepatitis C virus protease inhibitory  
 PT activity useful for treating disorders associated with hepatitis C  
 PT virus protease  
 XX  
 XX Claim 17; Page 65; 69pp; English.  
 XX  
 XX The sequence represents a peptide compound of the invention having  
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
 CC invention are alpha-ketoamide peptide analogues. The peptides have  
 CC virucide activity, and are useful for treating and in the manufacture of  
 CC a medicament to treat disorders associated with HCV protease. A  
 CC pharmaceutical composition comprising the peptide as an active ingredient  
 CC is useful for treating disorders associated with hepatitis C virus.  
 XX  
 XX Sequence 11 AA;

Query Match 85.2%; Score 46; DB 23; Length 11;  
 Best Local Similarity 90.9%; Pred. No. 0.022;  
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 EEVVPXGMDYS 11  
 |||||  
 DB 1 EEVVPXGTDYS 11

RESULT 14  
 ABB80548  
 ID ABB80548 standard; peptide; 11 AA.  
 XX  
 XX ABB80548;  
 XX  
 XX 08-OCT-2002 (first entry)  
 DT  
 XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #28.  
 DE  
 XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
 KW virucide.  
 KW  
 XX Synthetic.  
 OS  
 XX Key Location/Qualifiers  
 FH Modified-site 1 /note= "N-terminal acetyl"  
 FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with  
 residue 7"  
 FT Misc-difference 9 /note= "D-form residue"  
 FT Modified-site 11 /note= "C-terminal amide"  
 FT  
 XX

PN WO200208251-A2.  
PD 31-JAN-2002.  
XX 19-JUL-2001; 2001WO-US23169.  
XX 21-JUL-2000; 2000US-220101P.  
XX (CORV-) CORVAS INT INC.  
PA Lim-wilby M, Levy OE, Brunck TK;  
XX WPI; 2002-361643/39.  
XX Novel peptide compound having hepatitis C virus protease inhibitory activity useful for treating disorders associated with hepatitis C virus protease -  
XX Claim 17; Page 65; 69pp; English.  
XX The sequence represents a peptide compound of the invention having hepatitis C virus (HCV) protease inhibitory activity. The peptides of the invention are alpha-ketoamide peptide analogues. The peptides have virucide activity, and are useful for treating and in the manufacture of a medicament to treat disorders associated with HCV protease. A pharmaceutical composition comprising the peptide as an active ingredient is useful for treating disorders associated with hepatitis C virus.  
XX Query Match 85.2%; Score 46; DB 23; Length 11;  
XX Best Local Similarity 90.9%; Pred. No. 0.022;  
XX Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 EEVVPXGMDYS 11  
DB 1 EEVVPXGTDYS 11  
RESULT 15  
ABB0551  
ID ABB0551 standard; peptide; 11 AA.  
XX ABB0551;  
XX 08-OCT-2002 (first entry)  
XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #31.  
XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide; virucide.  
XX Synthetic.  
XX Key Location/Qualifiers  
FH Modified-site 1 /note= "N-terminal acetyl"  
FT Modified-site 6  
FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with residue 7"  
FT Misc-difference 9 /note= "D-form residue"  
FT Modified-site 11 /note= "C-terminal amide"  
XX WO200208251-A2.  
XX 31-JAN-2002.  
XX 19-JUL-2001; 2001WO-US23169.  
XX 21-JUL-2000; 2000US-220101P.

PA (CORV-) CORVAS INT INC.  
XX Lim-wilby M, Levy OE, Brunck TK;  
XX WPI; 2002-361643/39.  
XX Novel peptide compound having hepatitis C virus protease inhibitory activity useful for treating disorders associated with hepatitis C virus protease -  
XX Claim 17; Page 65; 69pp; English.  
XX The sequence represents a peptide compound of the invention having hepatitis C virus (HCV) protease inhibitory activity. The peptides of the invention are alpha-ketoamide peptide analogues. The peptides have virucide activity, and are useful for treating and in the manufacture of a medicament to treat disorders associated with HCV protease. A pharmaceutical composition comprising the peptide as an active ingredient is useful for treating disorders associated with hepatitis C virus.  
XX Query Match 85.2%; Score 46; DB 23; Length 11;  
XX Best Local Similarity 90.9%; Pred. No. 0.022;  
XX Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 EEVVPXGMDYS 11  
DB 1 EEVVPXGSDYS 11  
RESULT 16  
ABB0556  
ID ABB0556 standard; peptide; 11 AA.  
XX ABB0556;  
XX 08-OCT-2002 (first entry)  
XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #36.  
XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide; virucide.  
XX Synthetic.  
XX Key Location/Qualifiers  
FH Modified-site 1 /note= "N-terminal acetyl"  
FT Modified-site 6  
FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with residue 7"  
FT Misc-difference 8 /note= "D-form residue"  
FT Modified-site 11 /note= "C-terminal amide"  
XX WO200208251-A2.  
XX 31-JAN-2002.  
XX 19-JUL-2001; 2001WO-US23169.  
XX 21-JUL-2000; 2000US-220101P.  
XX (CORV-) CORVAS INT INC.  
XX Lim-wilby M, Levy OE, Brunck TK;  
XX WPI; 2002-361643/39.  
XX Novel peptide compound having hepatitis C virus protease inhibitory activity useful for treating disorders associated with hepatitis C



PT virus protease  
 XX Claim 17; Page 65; 69pp; English.  
 XX The sequence represents a peptide compound of the invention having  
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
 CC invention are alpha-ketoamide peptide analogues. The peptides have  
 CC virucide activity, and are useful for treating and in the manufacture of  
 CC a medicament to treat disorders associated with HCV protease. A  
 CC pharmaceutical composition comprising the peptide as an active ingredient  
 CC is useful for treating disorders associated with hepatitis C virus.  
 XX Sequence 11 AA;  
 SQ Query Match 85.2%; Score 46; DB 23; Length 11;  
 Best Local Similarity 90.9%; Pred. No. 0.022;  
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 EEVVPXGMDYS 11  
 Db 1 EEVVPXGSDYS 11  
 ||||| |||  
 RESULT 17  
 ABB80557  
 ID ABB80557 standard; peptide; 11 AA.  
 XX AC ABB80557;  
 XX 08-OCT-2002 (first entry)  
 DT Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #37.  
 DE Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
 KW virucide.  
 KW Synthetic.  
 OS Key Location/Qualifiers  
 FH Modified-site 1 /note= "N-terminal acetyl"  
 FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with  
 FT residue 7"  
 FT Misc-difference 8 /note= "D-form residue"  
 FT Misc-difference 9 /note= "D-form residue"  
 FT Modified-site 11 /note= "C-terminal amide"  
 FT WO200208251-A2.  
 PN 31-JAN-2002.  
 XX 19-JUL-2001; 2001WO-US23169.  
 XX 21-JUL-2000; 2000US-220101P.  
 XX (CORV-) CORVAS INT INC.  
 XX Lim-wilby M, Levy OE, Brunck TK;  
 XX WPI; 2002-361643/39.  
 XX Novel peptide compound having hepatitis C virus protease inhibitory  
 XX activity useful for treating disorders associated with hepatitis C  
 XX virus protease  
 XX Claim 17; Page 65; 69pp; English.  
 XX The sequence represents a peptide compound of the invention having  
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
 CC invention are alpha-ketoamide peptide analogues. The peptides have  
 CC virucide activity, and are useful for treating and in the manufacture of  
 CC a medicament to treat disorders associated with HCV protease. A  
 CC pharmaceutical composition comprising the peptide as an active ingredient  
 CC is useful for treating disorders associated with hepatitis C virus.

CC invention are alpha-ketoamide peptide analogues. The peptides have  
 CC virucide activity, and are useful for treating and in the manufacture of  
 CC a medicament to treat disorders associated with HCV protease. A  
 CC pharmaceutical composition comprising the peptide as an active ingredient  
 CC is useful for treating disorders associated with hepatitis C virus.  
 XX Sequence 11 AA;  
 SQ Query Match 85.2%; Score 46; DB 23; Length 11;  
 Best Local Similarity 90.9%; Pred. No. 0.022;  
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 EEVVPXGMDYS 11  
 Db 1 EEVVPXGSDYS 11  
 ||||| |||  
 RESULT 18  
 ABB80559  
 ID ABB80559 standard; peptide; 11 AA.  
 XX AC ABB80559;  
 XX 08-OCT-2002 (first entry)  
 DT Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #39.  
 DE Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
 KW virucide.  
 KW Synthetic.  
 OS Key Location/Qualifiers  
 FH Modified-site 1 /note= "N-terminal acetyl"  
 FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with  
 FT residue 7"  
 FT Misc-difference 8 /note= "D-form residue"  
 FT Misc-difference 8 /note= "D-form residue"  
 FT Modified-site 8 /note= "Oxymethionine"  
 FT Modified-site 11 /note= "C-terminal amide"  
 FT WO200208251-A2.  
 PN 31-JAN-2002.  
 XX 19-JUL-2001; 2001WO-US23169.  
 XX 21-JUL-2000; 2000US-220101P.  
 XX (CORV-) CORVAS INT INC.  
 XX Lim-wilby M, Levy OE, Brunck TK;  
 XX WPI; 2002-361643/39.  
 XX Novel peptide compound having hepatitis C virus protease inhibitory  
 XX activity useful for treating disorders associated with hepatitis C  
 XX virus protease  
 XX Claim 17; Page 65; 69pp; English.  
 XX The sequence represents a peptide compound of the invention having  
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
 CC invention are alpha-ketoamide peptide analogues. The peptides have  
 CC virucide activity, and are useful for treating and in the manufacture of  
 CC a medicament to treat disorders associated with HCV protease. A  
 CC pharmaceutical composition comprising the peptide as an active ingredient  
 CC is useful for treating disorders associated with hepatitis C virus.

SQ Sequence 11 AA;

Query Match 85.2%; Score 46; DB 23; Length 11;  
Best Local Similarity 90.9%; Pred. No. 0.022;  
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXGMDYS 11  
|||||  
Db 1 EEVVPXGMSYS 11

RESULT 19

ABB80563  
ID ABB80563 standard; peptide; 11 AA.

XX ABB80563;

XX 08-OCT-2002 (first entry)

XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #43.

XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;

KW virucide.

XX Synthetic.

XX OS

XX Key Location/Qualifiers

FH Modified-site 1

FT /note= "N-terminal acetyl"

FT Modified-site 6

FT /note= "Valyl carbonyl forming keto-amide linkage with residue 7"

FT Modified-site 11

FT /note= "C-terminal amide"

XX WO200208251-A2.

XX 31-JAN-2002.

XX 19-JUL-2001; 2001WO-US23169.

XX 21-JUL-2000; 2000US-220101P.

XX (CORV-) CORVAS INT INC.

XX Lim-wilby M, Levy OE, Brunck TK;

XX WPI; 2002-361643/39.

XX Novel peptide compound having hepatitis C virus protease inhibitory

XX activity useful for treating disorders associated with hepatitis C

XX virus protease

XX Claim 17; Page 65; 69pp; English.

XX The sequence represents a peptide compound of the invention having

XX hepatitis C virus (HCV) protease inhibitory activity. The peptides of the

XX invention are alpha-ketoamide peptide analogues. The peptides have

XX virucide activity, and are useful for treating and in the manufacture of

XX a medicament to treat disorders associated with HCV protease. A

XX pharmaceutical composition comprising the peptide as an active ingredient

XX is useful for treating disorders associated with hepatitis C virus.

XX SQ Sequence 11 AA;

XX Query Match 85.2%; Score 46; DB 23; Length 11;

XX Best Local Similarity 90.9%; Pred. No. 0.022;

XX Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

XX QY 1 EEVVPXGMDYS 11

XX |||||

XX Db 1 EEVVPXGMSYS 11

XX 08-OCT-2002 (first entry)

XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #45.

RESULT 20

ABB80564

XX ABB80564 standard; peptide; 11 AA.

XX ABB80564;

XX 08-OCT-2002 (first entry)

XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #44.

XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;

XX virucide.

XX Synthetic.

XX Key Location/Qualifiers

FH Modified-site 1

FT /note= "N-terminal acetyl"

FT Modified-site 6

FT /note= "Leucyl carbonyl forming keto-amide linkage with residue 7"

FT Modified-site 11

FT /note= "C-terminal amide"

XX WO200208251-A2.

XX 31-JAN-2002.

XX 19-JUL-2001; 2001WO-US23169.

XX 21-JUL-2000; 2000US-220101P.

XX (CORV-) CORVAS INT INC.

XX Lim-wilby M, Levy OE, Brunck TK;

XX WPI; 2002-361643/39.

XX Novel peptide compound having hepatitis C virus protease inhibitory

XX activity useful for treating disorders associated with hepatitis C

XX virus protease

XX Claim 17; Page 65; 69pp; English.

XX The sequence represents a peptide compound of the invention having

XX hepatitis C virus (HCV) protease inhibitory activity. The peptides of the

XX invention are alpha-ketoamide peptide analogues. The peptides have

XX virucide activity, and are useful for treating and in the manufacture of

XX a medicament to treat disorders associated with HCV protease. A

XX pharmaceutical composition comprising the peptide as an active ingredient

XX is useful for treating disorders associated with hepatitis C virus.

XX SQ Sequence 11 AA;

XX Query Match 85.2%; Score 46; DB 23; Length 11;

XX Best Local Similarity 90.9%; Pred. No. 0.022;

XX Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

XX QY 1 EEVVPXGMDYS 11

XX |||||

XX Db 1 EEVVPXGMSYS 11

XX 08-OCT-2002 (first entry)

XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #45.

	/note= "2-aminoisobutyl carbonyl residue forming a keto-amide linkage with residue 7"
Modified-site	11 /note= "C-terminal amide"
WO200208251-A2.	
31-JAN-2002.	
19-JUL-2001; 2001WO-US23169.	
21-JUL-2000; 2000US-220101P.	
(CORV-) CORVAS INT INC.	
Lim-wilby M, Levy OE, Brunck TK;	
WPI; 2002-361643/39.	
Novel peptide compound having hepatitis C virus protease inhibitory activity useful for treating disorders associated with hepatitis C virus protease -	
Claim 17; Page 65; 69pp; English.	
The sequence represents a peptide compound of the invention having hepatitis C virus (HCV) protease inhibitory activity. The peptides of the invention are alpha-ketoamide peptide analogues. The peptides have virucide activity, and are useful for treating and in the manufacture of a medicament to treat disorders associated with HCV protease. A pharmaceutical composition comprising the peptide as an active ingredient is useful for treating disorders associated with hepatitis C virus.	
Sequence	11 AA;
Query Match	85.2%; Score 46; DB 23; Length 11;
Best Local Similarity	90.9%; Pred. No. 0.022;
Matches	10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY	1 EEVVPXGMDYS 11 
Db	1 EEVVPXGMSYS 11 
RESULT 23	
ABB80567	
ID	ABB80567 standard; peptide; 11 AA.
XX AC	ABB80567;
XX DT	08-OCT-2002 (first entry)
XX DE	Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #47.
XX KW	Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide; virucide.
XX OS	Synthetic.
XX Key	Location/Qualifiers
FT Modified-site	1 /note= "N-terminal acetyl"
FT Modified-site	6 /note= "(S,s)allothreonyl carbonyl residue forming a keto-amide linkage with residue 7"
FT Modified-site	11 /note= "C-terminal amide"
WT	
PN WO200208251-A2.	
XX 31-JAN-2002.	
XX 19-JUL-2001; 2001WO-US23169.	
XX 21-JUL-2000; 2000US-220101P.	
XX (CORV-) CORVAS INT INC.	
XX Lim-wilby M, Levy OE, Brunck TK;	
XX WPI; 2002-361643/39.	
XX Novel peptide compound having hepatitis C virus protease inhibitory activity useful for treating disorders associated with hepatitis C virus protease -	
XX Claim 17; Page 65; 69pp; English.	
XX The sequence represents a peptide compound of the invention having hepatitis C virus (HCV) protease inhibitory activity. The peptides of the invention are alpha-ketoamide peptide analogues. The peptides have virucide activity, and are useful for treating and in the manufacture of a medicament to treat disorders associated with HCV protease. A pharmaceutical composition comprising the peptide as an active ingredient is useful for treating disorders associated with hepatitis C virus.	
XX Sequence	11 AA;
XX Query Match	85.2%; Score 46; DB 23; Length 11;
XX Best Local Similarity	90.9%; Pred. No. 0.022;
XX Matches	10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX QY	1 EEVVPXGMDYS 11 
XX Db	1 EEVVPXGMSYS 11 
XX RESULT 23	
XX ABB80567	
XX ID	ABB80567 standard; peptide; 11 AA.
XX AC	ABB80567;
XX DT	08-OCT-2002 (first entry)
XX DE	Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #47.
XX KW	Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide; virucide.
XX OS	Synthetic.
XX Key	Location/Qualifiers
XX FT Modified-site	1 /note= "N-terminal acetyl"
XX FT Modified-site	6 /note= "(S,s)allothreonyl carbonyl residue forming a keto-amide linkage with residue 7"
XX FT Modified-site	11 /note= "C-terminal amide"
XX WT	
XX PN WO200208251-A2.	
XX XX 31-JAN-2002.	
XX XX 19-JUL-2001; 2001WO-US23169.	
XX XX 21-JUL-2000; 2000US-220101P.	
XX XX (CORV-) CORVAS INT INC.	
XX XX Lim-wilby M, Levy OE, Brunck TK;	
XX XX WPI; 2002-361643/39.	
XX XX Novel peptide compound having hepatitis C virus protease inhibitory activity useful for treating disorders associated with hepatitis C virus protease -	
XX XX Claim 17; Page 65; 69pp; English.	
XX XX The sequence represents a peptide compound of the invention having hepatitis C virus (HCV) protease inhibitory activity. The peptides of the invention are alpha-ketoamide peptide analogues. The peptides have virucide activity, and are useful for treating and in the manufacture of a medicament to treat disorders associated with HCV protease. A pharmaceutical composition comprising the peptide as an active ingredient is useful for treating disorders associated with hepatitis C virus.	
XX XX Sequence	11 AA;
XX XX Query Match	85.2%; Score 46; DB 23; Length 11;
XX XX Best Local Similarity	90.9%; Pred. No. 0.022;
XX XX Matches	10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX XX QY	1 EEVVPXGMDYS 11 
XX XX Db	1 EEVVPXGMSYS 11 
XX XX RESULT 23	
XX XX ABB80567	
XX XX ID	ABB80567 standard; peptide; 11 AA.
XX XX AC	ABB80567;
XX XX DT	08-OCT-2002 (first entry)
XX XX DE	Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #47.
XX XX KW	Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide; virucide.
XX XX OS	Synthetic.
XX XX Key	Location/Qualifiers
XX XX FT Modified-site	1 /note= "N-terminal acetyl"
XX XX FT Modified-site	6 /note= "(S,s)allothreonyl carbonyl residue forming a keto-amide linkage with residue 7"
XX XX FT Modified-site	11 /note= "C-terminal amide"
XX XX WT	
XX XX PN WO200208251-A2.	
XX XX XX 31-JAN-2002.	
XX XX XX 19-JUL-2001; 2001WO-US23169.	
XX XX XX 21-JUL-2000; 2000US-220101P.	
XX XX XX (CORV-) CORVAS INT INC.	
XX XX XX Lim-wilby M, Levy OE, Brunck TK;	
XX XX XX WPI; 2002-361643/39.	
XX XX XX Novel peptide compound having hepatitis C virus protease inhibitory activity useful for treating disorders associated with hepatitis C virus protease -	
XX XX XX Claim 17; Page 65; 69pp; English.	
XX XX XX The sequence represents a peptide compound of the invention having hepatitis C virus (HCV) protease inhibitory activity. The peptides of the invention are alpha-ketoamide peptide analogues. The peptides have virucide activity, and are useful for treating and in the manufacture of a medicament to treat disorders associated with HCV protease. A pharmaceutical composition comprising the peptide as an active ingredient is useful for treating disorders associated with hepatitis C virus.	
XX XX XX Sequence	11 AA;
XX XX XX Query Match	85.2%; Score 46; DB 23; Length 11;
XX XX XX Best Local Similarity	90.9%; Pred. No. 0.022;
XX XX XX Matches	10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX XX XX QY	1 EEVVPXGMDYS 11 
XX XX XX Db	1 EEVVPXGMSYS 11 
XX XX XX RESULT 23	
XX XX XX ABB80567	
XX XX XX ID	ABB80567 standard; peptide; 11 AA.
XX XX XX AC	ABB80567;
XX XX XX DT	08-OCT-2002 (first entry)
XX XX XX DE	Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #47.
XX XX XX KW	Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide; virucide.
XX XX XX OS	Synthetic.
XX XX XX Key	Location/Qualifiers
XX XX XX FT Modified-site	1 /note= "N-terminal acetyl"
XX XX XX FT Modified-site	6 /note= "(S,s)allothreonyl carbonyl residue forming a keto-amide linkage with residue 7"
XX XX XX FT Modified-site	11 /note= "C-terminal amide"
XX XX XX WT	
XX XX XX PN WO200208251-A2.	
XX XX XX XX 31-JAN-2002.	
XX XX XX XX 19-JUL-2001; 2001WO-US23169.	
XX XX XX XX 21-JUL-2000; 2000US-220101P.	
XX XX XX XX (CORV-) CORVAS INT INC.	
XX XX XX XX Lim-wilby M, Levy OE, Brunck TK;	

XX 21-JUL-2000; 2000US-220101P.  
XX (CORV-) CORVAS INT INC.  
XX Lim-wilby M; Levy OE, Brunck TK;  
XX WPI; 2002-361643/39.  
XX Novel peptide compound having hepatitis C virus protease inhibitory  
PT activity useful for treating disorders associated with hepatitis C  
PT virus protease  
XX  
XX Claim 17; Page 65; 69pp; English.  
XX  
XX The sequence represents a peptide compound of the invention having  
CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
CC invention are alpha-ketoamide peptide analogues. The peptides have  
CC virucide activity, and are useful for treating and in the manufacture of  
CC a medicament to treat disorders associated with HCV protease. A  
CC pharmaceutical composition comprising the peptide as an active ingredient  
CC is useful for treating disorders associated with hepatitis C virus.  
XX  
XX Sequence 11 AA;  
XX

Query Match 85.2%; Score 45; DB 23; Length 11;  
Best Local Similarity 90.9%; Pred. NO. 0.022;  
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 EEVVPXGMDYS 11  
Db 1 EEVVPXGMSYS 11

RESULT 24  
ABB0568  
ID ABB0568 standard; peptide; 11 AA.  
XX AC ABB0568;  
XX  
XX 08-OCT-2002 (first entry)  
XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #48.  
XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
XX virucide.  
XX Synthetic.  
XX Key Location/Qualifiers  
FH Modified-site 1 /note= "N-terminal acetyl"  
FT Modified-site 6  
FT Modified-site 6 /note= "Alpha-propynyl-glycyl-carbonyl residue forming  
FT a keto-amide linkage with residue 7"  
FT Modified-site 11 /note= "C-terminal amide"  
FT  
XX WO200208251-A2.  
XX 31-JAN-2002.  
XX 19-JUL-2001; 2001WO-US23169.  
XX 21-JUL-2000; 2000US-220101P.  
XX (CORV-) CORVAS INT INC.  
XX Lim-wilby M; Levy OE, Brunck TK;  
XX WPI; 2002-361643/39.  
XX Novel peptide compound having hepatitis C virus protease inhibitory

PT activity useful for treating disorders associated with hepatitis C  
PT virus protease  
XX  
XX Claim 17; Page 65; 69pp; English.  
XX  
XX The sequence represents a peptide compound of the invention having  
CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
CC invention are alpha-ketoamide peptide analogues. The peptides have  
CC virucide activity, and are useful for treating and in the manufacture of  
CC a medicament to treat disorders associated with HCV protease. A  
CC pharmaceutical composition comprising the peptide as an active ingredient  
CC is useful for treating disorders associated with hepatitis C virus.  
XX  
XX Sequence 11 AA;  
XX

Query Match 85.2%; Score 46; DB 23; Length 11;  
Best Local Similarity 90.9%; Pred. NO. 0.022;  
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 EEVVPXGMDYS 11  
Db 1 EEVVPXGMSYS 11

RESULT 25  
ABB0523  
ID ABB0523 standard; peptide; 11 AA.  
XX AC ABB0523;  
XX  
XX 08-OCT-2002 (first entry)  
XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #3.  
XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
XX virucide.  
XX Synthetic.  
XX Key Location/Qualifiers  
FH Modified-site 1 /note= "N-terminal acetyl"  
FT Modified-site 6  
FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with  
FT residue 7"  
FT Misc-difference 9 /note= "D-form residue"  
FT Modified-site 11 /note= "C-terminal amide"  
FT  
XX WO200208251-A2.  
XX 31-JAN-2002.  
XX 19-JUL-2001; 2001WO-US23169.  
XX 21-JUL-2000; 2000US-220101P.  
XX (CORV-) CORVAS INT INC.  
XX Lim-wilby M; Levy OE, Brunck TK;  
XX WPI; 2002-361643/39.  
XX Novel peptide compound having hepatitis C virus protease inhibitory  
PT activity useful for treating disorders associated with hepatitis C  
PT virus protease  
XX  
XX Claim 17; Page 64; 69pp; English.  
XX  
XX The sequence represents a peptide compound of the invention having  
CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
CC invention are alpha-ketoamide peptide analogues. The peptides have

Wed Jun 11 15:44:44 2003

CC virucide activity, and are useful for treating and in the manufacture of  
 CC a medicament to treat disorders associated with HCV protease. A  
 CC pharmaceutical composition comprising the peptide as an active ingredient  
 CC is useful for treating disorders associated with hepatitis C virus.

XX  
 SQ Sequence 11 AA;

Query Match 83.3%; Score 45; DB 23; Length 11;  
 Best Local Similarity 90.9%; Pred. No. 0.034;  
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXGMDYS 11  
 |||||  
 Db 1 EEVVPXGMHYS 11

Search completed: June 10, 2003, 13:39:07  
 Job time : 32.3571 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2003, 13:31:45 ; Search time 9.64286 Seconds  
(without alignments)  
33.564 Million cell updates/sec

Title: US-09-909-164-8  
Perfect score: 54  
Sequence: 1 BEVXPXGMDYS 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues 262574  
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 75 summaries

Database : Issued Patents\_AA:\*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	36	66.7	341	4	US-08-853-948B-4
2	36	66.7	348	4	US-08-853-948B-5
3	36	66.7	368	4	US-09-697-367-24
4	36	66.7	1068	2	US-08-429-054A-11
5	36	66.7	1068	2	US-08-718-777-7
6	36	66.7	1068	3	US-09-051-341-7
7	34	63.0	140	4	US-08-569-147-76
8	34	63.0	140	4	US-08-569-147-82
9	33	61.1	59	4	US-08-963-851-14
10	33	61.1	378	1	US-08-070-165F-8
11	33	61.1	378	2	US-08-885-418-8
12	33	61.1	801	4	US-09-383-630-6
13	32	59.3	65	6	5177197-51
14	32	59.3	102	2	US-08-580-988A-23
15	32	59.3	152	2	US-08-460-694-4
16	32	59.3	152	3	US-08-460-744-4
17	32	59.3	152	3	US-07-667-711B-4
18	32	59.3	173	1	US-08-193-977-7
19	32	59.3	189	2	US-08-464-517-21
20	32	59.3	189	2	US-08-246-361A-21
21	32	59.3	189	3	US-08-463-772-21
22	32	59.3	189	5	PCT-US93-05000-21
23	32	59.3	231	3	US-08-926-842B-20
24	32	59.3	236	2	US-08-464-517-22
25	32	59.3	236	2	US-08-246-361A-22
26	32	59.3	236	3	US-08-463-772-22
27	32	59.3	236	5	PCT-US93-05000-22

28	32	59.3	240	3	US-08-926-842B-21
29	32	59.3	280	2	US-08-464-517-6
30	32	59.3	280	3	US-08-463-772-6
31	32	59.3	289	2	US-08-246-361A-4
32	32	59.3	289	5	PCT-US93-05000-4
33	32	59.3	291	5	PCT-US93-05000-6
34	32	59.3	292	2	US-08-464-517-23
35	32	59.3	292	2	US-08-246-361A-6
36	32	59.3	292	2	US-08-246-361A-23
37	32	59.3	292	3	US-08-463-772-23
38	32	59.3	292	5	PCT-US93-05000-23
39	32	59.3	295	1	US-07-947-120-8
40	32	59.3	295	1	US-08-472-893A-8
41	32	59.3	295	2	US-08-460-694-2
42	32	59.3	295	2	US-08-464-517-19
43	32	59.3	295	2	US-08-464-517-20
44	32	59.3	295	2	US-08-246-361A-19
45	32	59.3	295	2	US-08-246-361A-20
46	32	59.3	295	3	US-08-463-772-19
47	32	59.3	295	3	US-08-463-772-20
48	32	59.3	295	3	US-08-460-744-2
49	32	59.3	295	3	US-07-667-711B-2
50	32	59.3	295	3	US-08-947-492-8
51	32	59.3	295	5	PCT-US93-05000-2
52	32	59.3	295	5	PCT-US93-05000-19
53	32	59.3	295	5	PCT-US93-05000-20
54	32	59.3	309	2	US-08-464-517-4
55	32	59.3	309	3	US-08-463-772-4
56	32	59.3	410	6	5177197-1
57	32	59.3	493	4	US-09-411-628-10
58	32	59.3	618	2	US-08-770-761A-3
59	32	59.3	647	2	US-08-770-761A-8
60	32	59.3	660	2	US-08-770-761A-2
61	32	59.3	662	2	US-08-770-761A-5
62	32	59.3	705	2	US-08-770-761A-7
63	32	59.3	756	4	US-09-085-199B-9
64	32	59.3	819	2	US-08-464-517-7
65	32	59.3	819	2	US-08-246-361A-7
66	32	59.3	819	3	US-08-463-772-7
67	32	59.3	819	5	PCT-US93-05000-7
68	32	59.3	914	4	US-09-085-199B-4
69	32	59.3	1090	4	US-09-085-199B-5
70	32	59.3	1394	6	5177197-30
71	31	57.4	139	4	US-09-134-001C-5124
72	31	57.4	267	4	US-09-399-913-57
73	31	57.4	381	4	US-09-134-001C-3003
74	31	57.4	529	4	US-09-240-639-4
75	31	57.4	622	2	US-08-459-146-2

ALIGNMENTS

RESULT 1  
US-08-853-948B-4  
; Sequence 4, Application US/08853948B  
; Patent No. 6210943  
; GENERAL INFORMATION:  
; APPLICANT: AKIHAMA, Toyota  
; TITLE OF INVENTION: SUCCROSE PHOSPHATE SYNTHASE FROM CITRUS AND DNA ENCODING  
; TITLE OF INVENTION: THE SAME  
; FILE REFERENCE: 0049-0235-0  
; CURRENT APPLICATION NUMBER: US/08/853,948B  
; CURRENT FILING DATE: 1997-05-09  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 341  
; TYPE: PRT  
; ORGANISM: Citrus unshiu  
; FEATURE:  
; OTHER INFORMATION: Xaa at position 109 is one of Ala, Arg, Asn, Asp,  
; Cys, Gln, Glu, Gly, His, Ile, Leu, Lys, Met, Phe,  
;

OTHER INFORMATION: Pro, Ser, Thr, Trp, Tyr, or Val  
US-08-853-948B-4

Query Match 66.7%; Score 36; DB 4; Length 341;  
Best Local Similarity 66.7%; Pred. No. 14;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMDYS 11  
I:| | | | |  
Db 228 VIPPGMDFS 236

## RESULT 2

US-08-853-948B-5  
; Sequence 5, Application US/08853948B  
; Patent No. 6210943

GENERAL INFORMATION:  
; APPLICANT: AKIHAMA, Toyota

; TITLE OF INVENTION: SUCROSE PHOSPHATE SYNTHASE FROM CITRUS AND DNA ENCODING

; FILE REFERENCE: 0049-0235-0

; CURRENT APPLICATION NUMBER: US/08/853,948B

; CURRENT FILING DATE: 1997-05-09

; NUMBER OF SEQ ID NOS: 10

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 5

; LENGTH: 348

; TYPE: PRT

; ORGANISM: Citrus unshiu

US-08-853-948B-5

Query Match 66.7%; Score 36; DB 4; Length 348;  
Best Local Similarity 66.7%; Pred. No. 15;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMDYS 11  
I:| | | | |  
Db 234 VIPPGMDFS 242

## RESULT 3

US-09-697-367-24  
; Sequence 24, Application US/09697367  
; Patent No. 6323015

GENERAL INFORMATION:  
; APPLICANT: Orozco Jr., Emil M.

; APPLICANT: Caimi, Perry G.

; APPLICANT: Weng, Zude

; APPLICANT: Tarczynski, Mitchell

; TITLE OF INVENTION: SUCROSE PHOSPHATE SYNTHASE

; FILE REFERENCE: Bb1166 US NA

; CURRENT APPLICATION NUMBER: US/09/697,367

; CURRENT FILING DATE: 2000-10-26

; PRIOR APPLICATION NUMBER: 60/084,529

; PRIOR FILING DATE: 1998-MAY-07

; PRIOR APPLICATION NUMBER: PCT/US99/09865

; NUMBER OF SEQ ID NOS: 24

; SOFTWARE: Microsoft Office 97

; SEQ ID NO 24

; LENGTH: 368

; TYPE: PRT

; ORGANISM: Zea mays

US-09-697-367-24

Query Match 66.7%; Score 36; DB 4; Length 368;  
Best Local Similarity 66.7%; Pred. No. 16;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMDYS 11  
I:| | | | |  
Db 217 VIPPGMDFS 225

## RESULT 4

US-08-429-054A-11

; Sequence 11, Application US/08429054A

; Patent No. 5917126

GENERAL INFORMATION:

; APPLICANT: VAN ASSCHE, CHARLES; LANDO, DANIELLE; BRUNEAU,

; APPLICANT: JEAN, VOELKER, TONI; GERVAIS, MONICA

; TITLE OF INVENTION: SUCROSE PHOSPHATE SYNTHETASE (SPS),

; TITLE OF INVENTION: A PREPARATION METHOD AND CDNA THEREFOR, AND USE OF THE

; NUMBER OF SEQUENCES: 37

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: BIERMAN AND MUSERLIAN

; STREET: 600 THIRD AVENUE

; CITY: NEW YORK

; STATE: NEW YORK

; COUNTRY: USA

; ZIP: 10016

COMPUTER READABLE FORM:

; MEDIUM TYPE: FLOPPY DISK

; COMPUTER: IBM PC COMPATIBLE

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: WORDPERFECT 5.1

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/429,054A

; FILING DATE: 26-APR-1995

; CLASSIFICATION: 800

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 842,337

; FILING DATE: 20-March-1992

; APPLICATION NUMBER: PCT/FR 91/00593

; FILING DATE: 18-July-1991

; CLASSIFICATION: 800

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: French 90402094.9

; FILING DATE: 20-July-1990

; CLASSIFICATION: 800

; ATTORNEY/AGENT INFORMATION:

; NAME: Charles A. Muserlian

; REGISTRATION NUMBER: 19,683

; REFERENCE/DOCKET NUMBER: 146.1137

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 661-8000

; TELEFAX: (212) 661-8002

; INFORMATION FOR SEQ ID NO: 11:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1068

; TYPE: Amino acid

; STRANDEDNESS: Single

; TOPOLOGY: Unknown

; MOLECULE TYPE: Peptide

US-08-429-054A-11

Query Match 66.7%; Score 36; DB 2; Length 1068;  
Best Local Similarity 66.7%; Pred. No. 53;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMDYS 11  
I:| | | | |

Db 435 VIPPGMDFS 443

## RESULT 5

US-08-718-777-7

; Sequence 7, Application US/08718777

; Patent No. 5981852

GENERAL INFORMATION:

; APPLICANT: Van Assche, C.

; APPLICANT: Lando, D.

; APPLICANT: Bruneau, J. M.

; APPLICANT: Voelker, T.

; APPLICANT: Gervais, M.



;; TITLE OF INVENTION: MODIFICATION OF SUCROSE  
;; TITLE OF INVENTION: PHOSPHATE  
;; TITLE OF INVENTION: SYNTHASE IN PLANTS  
;; NUMBER OF SEQUENCES: 14  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Law Offices of Barbara Rae-Venter  
;; STREET: 260 Sheridan Avenue, Suite 440  
;; CITY: Palo Alto  
;; STATE: California  
;; COUNTRY: USA  
;; ZIP: 94306  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.30  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/718,777  
;; FILING DATE: NOT YET ASSIGNED  
;; CLASSIFICATION: 800  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/175,471  
;; FILING DATE: 27-DEC-1993  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Barbara Rae-Venter  
;; REGISTRATION NUMBER: 32,750  
;; REFERENCE/DOCKET NUMBER: CGNE.072.02US  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (415)328-4400  
;; TELEFAX: (415)328-4477  
;; INFORMATION FOR SEQ ID NO: 7:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 1068 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
US-08-718-777-7

Query Match 66.7%; Score 36; DB 2; Length 1068;  
Best Local Similarity 66.7%; Pred. No. 53;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMDYS 11  
|:| | | | |  
DB 435 VIPPGMDFS 443

RESULT 6  
US-09-051-341-7  
; Sequence 7, Application US/09051341  
; Patent No. 6124528  
; GENERAL INFORMATION:  
; APPLICANT: Shewmaker, C. K.  
; TITLE OF INVENTION: MODIFICATION OF SOLUBLE SOLIDS USING  
; TITLE OF INVENTION: SUCROSE PHOSPHATE SYNTHASE ENCODING SEQUENCE  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Rae-Venter Law Group, P.C.  
; STREET: 260 Sheridan Avenue, Suite 440  
; CITY: Palo Alto  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94306  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/051,341  
; FILING DATE:  
; CLASSIFICATION: 800  
; PRIOR APPLICATION DATA:

;; APPLICATION NUMBER: PCT/US96/17351  
;; FILING DATE: 25-OCT-1996  
;; APPLICATION NUMBER: US 08/549,016  
;; FILING DATE: 27-OCT-1995  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/372,200  
;; FILING DATE: 12-JAN-1995  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Barbara Rae-Venter, Ph.D.,  
;; REGISTRATION NUMBER: 32,750  
;; REFERENCE/DOCKET NUMBER: CGNE.110.02US  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (415)328-4400  
;; TELEFAX: (415)328-4477  
;; INFORMATION FOR SEQ ID NO: 7:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 1068 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
US-09-051-341-7

Query Match 66.7%; Score 36; DB 3; Length 1068;  
Best Local Similarity 66.7%; Pred. No. 53;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMDYS 11  
|:| | | | |  
DB 435 VIPPGMDFS 443

RESULT 7  
US-08-569-147-76  
; Sequence 76, Application US/08569147  
; Patent No. 6180377  
; GENERAL INFORMATION:  
; APPLICANT:  
; TITLE OF INVENTION: HUMANISED ANTIBODIES  
; NUMBER OF SEQUENCES: 95  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &  
; STREET: One Liberty Place - 46th Floor  
; CITY: Philadelphia  
; STATE: PA  
; COUNTRY: U.S.A.  
; ZIP: 19103  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)  
; CURRENT APPLICATION DATA: US/08/569,147  
; APPLICATION NUMBER: US/08/569,147  
; FILING DATE: 25-March-1996  
; CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Trujillo, Doreen Yanko  
; REGISTRATION NUMBER: 35,719  
; REFERENCE/DOCKET NUMBER: CARP-0047  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (215) 568-3100  
; TELEFAX: (215) 568-3439  
; INFORMATION FOR SEQ ID NO: 76:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 140 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-569-147-76

Query Match 63.0%; Score 34; DB 4; Length 140;  
Best Local Similarity 75.0%; Pred. No. 13;

Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 3 VVPXGMDY 10  
Db 122 VVPTGFDY 129

RESULT 8  
US-08-569-147-82  
; Sequence 82, Application US/08569147  
; Patent No. 6180377  
; GENERAL INFORMATION:  
; APPLICANT: HUMANISED ANTIBODIES  
; TITLE OF INVENTION: HUMANISED ANTIBODIES  
; NUMBER OF SEQUENCES: 95  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &  
; ADDRESSEE: No. 6180377ris, LLP  
; STREET: One Liberty place - 46th Floor  
; CITY: Philadelphia  
; STATE: PA  
; COUNTRY: U.S.A.  
; ZIP: 19103  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/569,147  
; FILING DATE: 25-March-1996  
; CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Trujillo, Doreen Yanko  
; REGISTRATION NUMBER: 35,719  
; REFERENCE/DOCKET NUMBER: CARP-0047  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (215) 568-3100  
; TELEFAX: (215) 568-3439  
; INFORMATION FOR SEQ ID NO: 82:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 140 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-569-147-82

Query Match 63.0%; Score 34; DB 4; Length 140;  
Best Local Similarity 75.0%; Pred. No. 13;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 VVPXGMDY 10  
Db 122 VVPTGFDY 129

RESULT 9  
US-08-963-851-14  
; Sequence 14, Application US/08963851  
; Patent No. 6300116  
; GENERAL INFORMATION:  
; APPLICANT: VAN DER OSTEN, CLAUS  
; APPLICANT: HALKIER, TORDEN  
; APPLICANT: ANDERSEN, CARSTEN  
; APPLICANT: BAUDITZ, PETER  
; APPLICANT: HANSEN, PETER KAMP  
; TITLE OF INVENTION: PROTEASE VARIANTS AND COMPOSITIONS  
; FILE REFERENCE: 4946,200-US  
; CURRENT APPLICATION NUMBER: US/08/963,851  
; CURRENT FILING DATE: 1997-11-04  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 14

; LENGTH: 59  
; TYPE: PRT  
; ORGANISM: Enterococcus faecalis  
US-08-963-851-14

Query Match 61.1%; Score 33; DB 4; Length 59;  
Best Local Similarity 45.5%; Pred. No. 7.5;  
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;  
QY 1 BEVVPXGMDYS 11  
Db 38 EKHIPGGLEYS 48

RESULT 10  
US-08-070-165F-8  
; Sequence 8, Application US/08070165F  
; Patent No. 5750365  
; GENERAL INFORMATION:  
; APPLICANT: Chiu, Ing-Ming  
; APPLICANT: Poulin, Matthew L  
; TITLE OF INVENTION: Acidic Fibroblast Growth Factor (aFGF)  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Ing-Ming Chiu  
; STREET: S2052 Davis Medical Research Center, 480 West  
; STREET: 9th Avenue  
; CITY: Columbus  
; STATE: Ohio  
; COUNTRY: USA  
; ZIP: 43210  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/070,165F  
; FILING DATE:  
; CLASSIFICATION: 435  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (614)-293-8093  
; TELEFAX: (614)-293-5631  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 378 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-070-165F-8

Query Match 61.1%; Score 33; DB 1; Length 378;  
Best Local Similarity 85.7%; Pred. No. 64;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 PXGMDYS 11  
Db 139 PFGMDYS 145

RESULT 11  
US-08-885-418-8  
; Sequence 8, Application US/08885418  
; Patent No. 5925528  
; GENERAL INFORMATION:  
; APPLICANT: Chiu, Ing-Ming  
; APPLICANT: Poulin, Matthew L  
; TITLE OF INVENTION: Acidic Fibroblast Growth Factor (aFGF)  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Ing-Ming Chiu  
; STREET: S2052 Davis Medical Research Center, 480 West  
; STREET: 9th Avenue

CITY: Columbus  
STATE: Ohio  
COUNTRY: USA  
ZIP: 43210  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/885,418  
FILING DATE:  
CLASSIFICATION: 435  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (614)-293-8093  
TELEFAX: (614)-293-5631  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 378 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-885-418-8

Query Match  
Best Local Similarity 61.1%; Score 33; DB 2; Length 378;  
Best Local Similarity 85.7%; Pred. No. 64;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 PXGMDYS 11  
Db 139 PFGMDYS 145

RESULT 12  
US-09-383-630-6  
Sequence 6, Application US/09383630A  
Patent No. 6265632  
GENERAL INFORMATION:  
APPLICANT: Avner Yayon et al.  
TITLE OF INVENTION: ANIMAL MODEL FOR FIBROBLAST GROWTH  
FACTOR RECEPTOR ASSOCIATED  
CHONDRODYSPLASIA  
NUMBER OF SEQUENCES: 18  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Mark M. Friedman c/o Anthony Castorina  
STREET: 2001 Jefferson Davis Highway, Suite 207  
CITY: Arlington  
STATE: Virginia  
COUNTRY: United States of America  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 1.44 megabyte, 3.5" microdisk  
COMPUTER: Twinhead\* Slimnote-890TX  
OPERATING SYSTEM: MS DOS version 6.2,  
Windows version 3.11  
SOFTWARE: Word for Windows version 2.0 converted  
to an ASCII file  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/383,630A  
FILING DATE: 26-Aug-1999  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: <Unknown>  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Friedman, Mark M.  
REGISTRATION NUMBER: 33,883  
REFERENCE/DOCKET NUMBER: 1402/2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 972-3-562553  
TELEFAX: 972-3-562554  
INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:  
LENGTH: 801  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 6:  
US-09-383-630-6

Query Match  
Best Local Similarity 61.1%; Score 33; DB 4; Length 801;  
Best Local Similarity 85.7%; Pred. No. 1.5e+02;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 PXGMDYS 11  
Db 566 PFGMDYS 572

RESULT 13  
5177197-51  
Patent No. 5177197  
APPLICANT: KANZAKI, TETSUTO; OLOFSSON, ANDERS; MOREN, ANITA;  
WERNSTEDT, CHRISTER; HELLMAN, ULF; MIYAZONO, KOHEI; CLAESSON-WELSH,  
LENA; HELDIN, CARL-HENRIK  
TITLE OF INVENTION: ISOLATED NUCLEOTIDE SEQUENCE EXPRESSING  
HUMAN TRANSFORMING GROWTH FACTOR-BETAL-BINDING PROTEIN  
NUMBER OF SEQUENCES: 53  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/487,343  
FILING DATE: 27-FEB-1990  
SEQ ID NO: 51:  
LENGTH: 65  
5177197-51

Query Match  
Best Local Similarity 59.3%; Score 32; DB 6; Length 65;  
Best Local Similarity 45.5%; Pred. No. 13;  
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMDYS 11  
Db 52 KEICPGMGYT 52

RESULT 14  
US-08-580-988A-23  
Sequence 23, Application US/08580988A  
Patent No. 5856161  
GENERAL INFORMATION:  
APPLICANT: Aggarwal et al.  
TITLE OF INVENTION: Tumor Necrosis Factor  
RECEPTOR-I-Associated Protein Kinase And Methods  
FOR ITS USE  
TITLE OF INVENTION: For Its Use  
NUMBER OF SEQUENCES: 27  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dr. Benjamin A. Adler  
STREET: 8011 Candle Lane  
CITY: Houston  
STATE: Texas  
COUNTRY: USA  
ZIP: 77071  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 1.44 Mb floppy disk  
COMPUTER: Apple Macintosh  
OPERATING SYSTEM: Macintosh  
SOFTWARE: Microsoft Word for Macintosh  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/580,988A  
FILING DATE: January 3, 1996  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Benjamin Aaron Adler, Ph.D., J.D.

REGISTRATION NUMBER: 35,423  
REFERENCE/DOCKET NUMBER: D5721CIP2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 713-777-2321  
TELEFAX: 713-777-6908  
INFORMATION FOR SEQ ID NO: 23:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 102 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
DESCRIPTION: no  
HYPOTHETICAL: no  
ANTI-SENSE: no  
FRAGMENT TYPE: internal  
ORIGINAL SOURCE:  
US-08-580-988A-23

Query Match 59.3%; Score 32; DB 2; Length 102;  
Best Local Similarity 60.0%; Pred. No. 22;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 EEVVPXGMDY 10  
Db 24 EEVFPAMNY 33

RESULT 15  
US-08-460-694-4  
Sequence 4, Application US/08460694  
Patent No. 5858655

GENERAL INFORMATION:  
APPLICANT: Arnold, Andrew  
TITLE OF INVENTION: Pradi Cyclin and its CDNA  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
STREET: 1100 New York Avenue, N.W., Suite 600  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/460,694  
FILING DATE: 02-JUN-1995  
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:  
NAME: McConathy, Evelyn H.  
REGISTRATION NUMBER: 35,279  
REFERENCE/DOCKET NUMBER: 0609.4070002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-2600  
TELEFAX: 202-371-2540

INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 152 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-460-694-4

Query Match 59.3%; Score 32; DB 2; Length 152;  
Best Local Similarity 60.0%; Pred. No. 35;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 EEVVPXGMDY 10

Db 20 EEVFPAMNY 29

RESULT 16  
US-08-460-744-4  
Sequence 4, Application US/08460744  
Patent No. 6107541

GENERAL INFORMATION:  
APPLICANT: Arnold, Andrew  
TITLE OF INVENTION: Pradi Cyclin and its CDNA  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
STREET: 1100 New York Avenue, N.W., Suite 600  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/460,744  
FILING DATE: 02-JUN-1995  
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:  
NAME: McConathy, Evelyn H.  
REGISTRATION NUMBER: 35,279  
REFERENCE/DOCKET NUMBER: 0609.4070005  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-2600  
TELEFAX: 202-371-2540

INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 152 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-460-744-4

Query Match 59.3%; Score 32; DB 3; Length 152;  
Best Local Similarity 60.0%; Pred. No. 35;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 EEVVPXGMDY 10  
Db 20 EEVFPAMNY 29

RESULT 17  
US-07-667-711B-4  
Sequence 4, Application US/07667711B  
Patent No. 6110700

GENERAL INFORMATION:  
APPLICANT: ARNOLD, ANDREW  
TITLE OF INVENTION: Pradi Cyclin and its CDNA  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
STREET: 1100 NEW YORK AVE., NW, SUITE 600  
CITY: WASHINGTON  
STATE: DC  
COUNTRY: USA  
ZIP: 20005

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30

us-09-909-164-8.ra1

Wed Jun 11 15:44:47 2003

;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/07/667,711B  
;; FILING DATE: 11-MAR-1991  
;; CLASSIFICATION: 435  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: MCPHAIL, DONALD R.  
;; REGISTRATION NUMBER: 35,811  
;; REFERENCE/DOCKET NUMBER: 0609.4070000  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (202) 371-2600  
;; TELEFAX: (202) 371-2540  
;; INFORMATION FOR SEQ ID NO: 4:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 152 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: not relevant  
;; MOLECULE TYPE: peptide  
US-07-667-711B-4

Query Match 59.3%; Score 32; DB 3; Length 152;  
Best Local Similarity 60.0%; Pred. No. 35;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Oy 1 EEVVPXGMDY 10  
Db 20 EEVFPPLAMNY 29

RESULT 18  
US-08-193-977-7  
; Sequence 7, Application US/08193977  
; Patent No. 5625031  
; GENERAL INFORMATION:  
; APPLICANT: WEBSTER, KEVIN R.  
; APPLICANT: COLEMAN, KEVIN G.  
; TITLE OF INVENTION: PEPTIDE INHIBITORS OF THE P33CDK2 AND  
; TITLE OF INVENTION: P34CDC2 CELL CYCLE REGULATORY KINASES AND HUMAN  
; TITLE OF INVENTION: PAPILLOMAVIRUS E7 ONCOPROTEIN  
; NUMBER OF SEQUENCES: 34  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: REED & ROBINS  
; STREET: 635 BRYANT STREET  
; CITY: PALO ALTO  
; STATE: CALIFORNIA  
; COUNTRY: UNITED STATES OF AMERICA  
; ZIP: 94301  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/193,977  
; FILING DATE: 08-FEB-1994  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: ROBINS, ROBERTA L.  
; REGISTRATION NUMBER: 33,208  
; REFERENCE/DOCKET NUMBER: 5998-0016  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 617-8999  
; TELEFAX: (415) 327-3231  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 173 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-193-977-7

Query Match 59.3%; Score 32; DB 1; Length 173;

Best Local Similarity 60.0%; Pred. No. 41;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Oy 1 EEVVPXGMDY 10  
Db 55 EEVFPPLAMNY 64

RESULT 19  
US-08-464-517-21  
; Sequence 21, Application US/08464517  
; Patent No. 5869640  
; GENERAL INFORMATION:  
; APPLICANT: BEACH, DAVID H.  
; TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO  
; NUMBER OF SEQUENCES: 50  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD  
; STREET: 60 State Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII(text)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/464,517  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/963,308  
; FILING DATE: 16-OCT-1992  
; APPLICATION NUMBER: US 07/888,178  
; FILING DATE: 26-MAY-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/701,514  
; FILING DATE: 16-MAY-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Matthew P. Vincent  
; REGISTRATION NUMBER: 36,709  
; REFERENCE/DOCKET NUMBER: MII-004C  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 227-7400  
; TELEFAX: (617) 227-5941  
; INFORMATION FOR SEQ ID NO: 21:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 189 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-464-517-21

Query Match 59.3%; Score 32; DB 2; Length 189;  
Best Local Similarity 60.0%; Pred. No. 46;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Oy 1 EEVVPXGMDY 10  
Db 74 EEVFPPLAMNY 83

RESULT 20  
US-08-246-361A-21  
; Sequence 21, Application US/08246361A  
; Patent No. 5998582  
; GENERAL INFORMATION:  
; APPLICANT: BEACH, DAVID H.  
; TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO  
; NUMBER OF SEQUENCES: 50

```

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII(text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/246,361A
; FILING DATE: 19-MAY-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/963,308
; FILING DATE: 16-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/888,178
; FILING DATE: 26-MAY-1992
; APPLICATION NUMBER: US 07/701,514
; FILING DATE: 16-MAY-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Matthew P. Vincent
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: MII-004C
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 189 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-246-361A-21

```

```

Query Match          59.3%; Score 32; DB 2; Length 189;
Best Local Similarity 60.0%; Pred. No. 46;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMDY 10
   ||| | | |
Db 74 EEVFPPLAMNY 83

```

```

RESULT 21
US-08-463-772-21
; Sequence 21, Application US/08463772
; Patent No. 6066501
; GENERAL INFORMATION:
; APPLICANT: BEACH, David H.
; TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII(text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/463,772
; FILING DATE:

```

```

; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/963,308
; FILING DATE: 16-OCT-1992
; APPLICATION NUMBER: US 07/888,178
; FILING DATE: 26-MAY-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/701,514
; FILING DATE: 16-MAY-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Matthew P. Vincent
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: MII-004C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 189 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-463-772-21

```

```

Query Match          59.3%; Score 32; DB 3; Length 189;
Best Local Similarity 60.0%; Pred. No. 46;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMDY 10
   ||| | | |
Db 74 EEVFPPLAMNY 83

```

```

RESULT 22
PCT-US93-05000-21
; Sequence 21, Application PC/TUS9305000
; GENERAL INFORMATION:
; APPLICANT: MITOTIX
; TITLE OF INVENTION: D-Type Cyclin and Uses Related Thereto
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: US
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/05000
; FILING DATE: 19930525
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/888,178
; FILING DATE: 26-MAY-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: CSHL91-02A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-6240
; TELEFAX: 616-861-9540
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 189 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: unknown
; MOLECULE TYPE: protein

```

PCT-US93-05000-21

Query Match 59.3%; Score 32; DB 5; Length 189;  
Best Local Similarity 60.0%; Pred. No. 46;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMDY 10  
||| | | |  
Db 74 EEVFLAMNY 83

## RESULT 23

US-08-926-842B-20  
; Sequence 20, Application US/08926842B  
; Patent No. 6030807  
; GENERAL INFORMATION:  
; APPLICANT: Sa-No. 6030807ueira, Isabel  
; APPLICANT: de Lencastre, Herminia  
; TITLE OF INVENTION: HIGHLY REGULABLE PROMOTER FOR HETEROLOGOUS GENE  
; TITLE OF INVENTION: EXPRESSION  
; NUMBER OF SEQUENCES: 64  
; CORRESPONDENCE ADDRESS:  
; ADDRESSER: Klauber & Jackson  
; STREET: 411 Hackensack Avenue  
; CITY: Hackensack  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/926,842B  
FILING DATE: 10-SEP-1997  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Jackson Esq., David A.  
REGISTRATION NUMBER: 26,742  
REFERENCE/DOCKET NUMBER: 600-1-089 N  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201 487-5800  
TELEFAX: 201 343-1684  
TELEX: 133521

## INFORMATION FOR SEQ ID NO: 20:

SEQUENCE CHARACTERISTICS:  
LENGTH: 231 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
ORGANISM: Escherichia coli  
US-08-926-842B-20

Query Match 59.3%; Score 32; DB 3; Length 231;  
Best Local Similarity 55.6%; Pred. No. 58;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 VVPXGMDYS 11  
: | | | | |  
Db 41 IKPSGVDYS 49

## RESULT 24

US-08-464-517-22  
; Sequence 22, Application US/08464517  
; Patent No. 5869640  
; GENERAL INFORMATION:  
; APPLICANT: BEACH, David H.  
; TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO

NUMBER OF SEQUENCES: 50  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD  
STREET: 60 State Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII(text)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/464,517  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/963,308  
FILING DATE: 16-OCT-1992  
APPLICATION NUMBER: US 07/888,178  
FILING DATE: 26-MAY-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/701,514  
FILING DATE: 16-MAY-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Matthew P. Vincent  
REGISTRATION NUMBER: 36,709  
REFERENCE/DOCKET NUMBER: MII-004C  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 227-7400  
TELEFAX: (617) 227-5941  
INFORMATION FOR SEQ ID NO: 22:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 236 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-464-517-22

Query Match 59.3%; Score 32; DB 2; Length 236;  
Best Local Similarity 60.0%; Pred. No. 59;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMDY 10  
||| | | |  
Db 20 EEVFLAMNY 29

## RESULT 25

US-08-246-361A-22  
; Sequence 22, Application US/08246361A  
; Patent No. 5998582  
; GENERAL INFORMATION:  
; APPLICANT: BEACH, David H.  
; TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO  
; NUMBER OF SEQUENCES: 50  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD  
; STREET: 60 State Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII(text)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/246,361A  
FILING DATE: 19-MAY-1994

CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/963,308  
FILING DATE: 16-OCT-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/888,178  
FILING DATE: 26-MAY-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/701,514  
FILING DATE: 16-MAY-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Matthew P. Vincent  
REGISTRATION NUMBER: 36,709  
REFERENCE/DOCKET NUMBER: MII-004C  
TELEPHONE: (617) 227-7400  
TELEFAX: (617) 227-5941  
INFORMATION FOR SEQ ID NO: 22:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 236 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-246-361A-22

Query Match 59.3%; Score 32; DB 2; Length 236;  
Best Local Similarity 60.0%; Pred. No. 59;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMDY 10  
||| | | |  
Db 20 EEVFPFLAMNY 29

Search completed: June 10, 2003, 13:51:32  
Job time : 10.6429 secs



20	33	61.1	736	9	US-10-176-758-420	Sequence 420, App
21	33	61.1	736	9	US-10-175-737-420	Sequence 420, App
22	33	61.1	736	9	US-10-173-706-420	Sequence 420, App
23	33	61.1	736	9	US-10-175-738-420	Sequence 420, App
24	33	61.1	736	9	US-10-175-752-420	Sequence 420, App
25	33	61.1	736	9	US-10-176-482-420	Sequence 420, App
26	33	61.1	736	9	US-10-176-757-420	Sequence 420, App
27	33	61.1	736	9	US-10-176-913-420	Sequence 420, App
28	33	61.1	736	9	US-10-180-552-420	Sequence 420, App
29	33	61.1	736	9	US-10-180-557-420	Sequence 420, App
30	33	61.1	736	9	US-10-173-700-420	Sequence 420, App
31	33	61.1	736	9	US-10-174-572-420	Sequence 420, App
32	33	61.1	736	9	US-10-174-579-420	Sequence 420, App
33	33	61.1	736	9	US-10-174-582-420	Sequence 420, App
34	33	61.1	736	9	US-10-174-588-420	Sequence 420, App
35	33	61.1	736	9	US-10-175-739-420	Sequence 420, App
36	33	61.1	736	9	US-10-175-740-420	Sequence 420, App
37	33	61.1	736	9	US-10-175-743-420	Sequence 420, App
38	33	61.1	736	9	US-10-176-488-420	Sequence 420, App
39	33	61.1	736	9	US-10-176-492-420	Sequence 420, App
40	33	61.1	736	9	US-10-176-747-420	Sequence 420, App
41	33	61.1	736	9	US-10-176-750-420	Sequence 420, App
42	33	61.1	736	9	US-10-176-985-420	Sequence 420, App
43	33	61.1	736	9	US-10-176-991-420	Sequence 420, App
44	33	61.1	736	9	US-10-176-992-420	Sequence 420, App
45	33	61.1	736	9	US-10-176-993-420	Sequence 420, App
46	33	61.1	736	9	US-10-184-658-420	Sequence 420, App
47	33	61.1	736	9	US-10-173-695-420	Sequence 420, App
48	33	61.1	736	9	US-10-173-697-420	Sequence 420, App
49	33	61.1	736	9	US-10-173-697-420	Sequence 420, App
50	33	61.1	736	9	US-10-173-705-420	Sequence 420, App
51	33	61.1	736	9	US-10-174-576-420	Sequence 420, App
52	33	61.1	736	9	US-10-174-585-420	Sequence 420, App
53	33	61.1	736	9	US-10-174-586-420	Sequence 420, App
54	33	61.1	736	9	US-10-175-747-420	Sequence 420, App
55	33	61.1	736	9	US-10-176-481-420	Sequence 420, App
56	33	61.1	736	9	US-10-176-485-420	Sequence 420, App
57	33	61.1	736	9	US-10-176-487-420	Sequence 420, App
58	33	61.1	736	9	US-10-176-493-420	Sequence 420, App
59	33	61.1	736	9	US-10-176-756-420	Sequence 420, App
60	33	61.1	736	9	US-10-176-911-420	Sequence 420, App
61	33	61.1	736	9	US-10-176-919-420	Sequence 420, App
62	33	61.1	736	9	US-10-176-925-420	Sequence 420, App
63	33	61.1	736	9	US-10-176-978-420	Sequence 420, App
64	33	61.1	736	9	US-10-176-978-420	Sequence 420, App
65	33	61.1	736	9	US-10-179-510-420	Sequence 420, App
66	33	61.1	736	9	US-10-180-543-420	Sequence 420, App
67	33	61.1	736	9	US-10-180-544-420	Sequence 420, App
68	33	61.1	736	9	US-10-180-546-420	Sequence 420, App
69	33	61.1	736	9	US-10-180-547-420	Sequence 420, App
70	33	61.1	736	9	US-10-180-549-420	Sequence 420, App
71	33	61.1	736	9	US-10-180-553-420	Sequence 420, App
72	33	61.1	736	9	US-10-181-000-420	Sequence 420, App
73	33	61.1	736	9	US-10-183-010-420	Sequence 420, App
74	33	61.1	736	9	US-10-183-013-420	Sequence 420, App
75	33	61.1	736	9	US-10-184-614-420	Sequence 420, App

ALIGNMENTS

RESULT 1  
US-10-217-700-10  
; Sequence 10, Application US/10217700  
; Publication NO. US20030070191A1  
; GENERAL INFORMATION:  
; APPLICANT: Haigler, Candace H.  
; APPLICANT: Holaday, A. Scott  
; TITLE OF INVENTION: TRANSGENIC FIBER PRODUCING PLANTS WITH INCREASED  
; TITLE OF INVENTION: EXPRESSION OF SUCROSE PHOSPHATE SYNTHASE  
; FILE REFERENCE: 201304/1000  
; CURRENT APPLICATION NUMBER: US/10/217,700  
; CURRENT FILING DATE: 2002-08-12

OM protein - protein search, using sw model  
Run on: June 10, 2003, 13:46:50 ; Search time 15 Seconds  
(without alignments)  
75.710 Million cell updates/sec

Title: US-09-909-164-8  
Perfect score: 54  
Sequence: 1 BEVVPXGMDYS 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 392085 seqs, 103240269 residues

Total number of hits satisfying chosen parameters: 392085

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 75 summaries

Database : Published Applications AA: \*  
1: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB pep: \*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB pep: \*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB pep: \*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB pep: \*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB pep: \*  
6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB pep: \*  
7: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB pep: \*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB pep: \*  
9: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB pep: \*  
10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB pep: \*  
11: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB pep: \*  
12: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB pep: \*  
13: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB pep: \*  
14: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB pep: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	36	66.7	1049	9	US-10-217-700-10
2	36	66.7	1068	9	US-10-217-700-8
3	36	66.7	1081	9	US-10-217-700-4
4	36	66.7	1083	9	US-10-217-700-11
5	36	66.7	1084	9	US-10-217-700-9
6	35	64.8	440	9	US-09-813-408-27
7	34	63.0	1062	10	US-09-815-242-5111
8	34	63.0	3472	9	US-10-027-806-4
9	34	63.0	3472	9	US-10-034-623-4
10	34	63.0	3472	9	US-10-027-801-4
11	33	61.1	59	10	US-09-948-080-14
12	33	61.1	283	9	US-09-738-626-4881
13	33	61.1	299	10	US-09-815-242-10697
14	33	61.1	736	9	US-09-978-295A-526
15	33	61.1	736	9	US-09-978-697-526
16	33	61.1	736	9	US-09-978-192A-526
17	33	61.1	736	9	US-09-999-832A-526
18	33	61.1	736	9	US-09-978-189-526
19	33	61.1	736	9	US-10-174-590-420

; EARLIER APPLICATION NUMBER: 09/394,272  
; EARLIER FILING DATE: 1999-09-10  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 10  
; LENGTH: 1049  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
US-10-217-700-10

Query Match  
Best Local Similarity 66.7%; Score 36; DB 9; Length 1049;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMDYS 11  
I:| | | | |  
Db 436 VIPPGMDFS 444

## RESULT 2

US-10-217-700-8  
; Sequence 8, Application US/10217700  
; Publication No. US20030070191A1  
; GENERAL INFORMATION:  
; APPLICANT: Haigler, Candace H.  
; TITLE OF INVENTION: TRANSGENIC FIBER PRODUCING PLANTS WITH INCREASED  
; FILE REFERENCE: 201304/1000  
; CURRENT APPLICATION NUMBER: US/10/217,700  
; EARLIER FILING DATE: 2002-08-12  
; EARLIER APPLICATION NUMBER: 09/394,272  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 8  
; LENGTH: 1068  
; TYPE: PRT  
; ORGANISM: Zea mays  
US-10-217-700-8

Query Match  
Best Local Similarity 66.7%; Score 36; DB 9; Length 1068;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMDYS 11  
I:| | | | |  
Db 435 VIPPGMDFS 443

## RESULT 3

US-10-217-700-4  
; Sequence 4, Application US/10217700  
; Publication No. US20030070191A1  
; GENERAL INFORMATION:  
; APPLICANT: Haigler, Candace H.  
; TITLE OF INVENTION: TRANSGENIC FIBER PRODUCING PLANTS WITH INCREASED  
; FILE REFERENCE: 201304/1000  
; CURRENT APPLICATION NUMBER: US/10/217,700  
; EARLIER FILING DATE: 2002-08-12  
; EARLIER APPLICATION NUMBER: 09/394,272  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 4  
; LENGTH: 1081  
; TYPE: PRT  
; ORGANISM: Craterostigma plantaginenum  
US-10-217-700-4

Query Match  
Best Local Similarity 66.7%; Score 36; DB 9; Length 1081;

Best Local Similarity 66.7%; Pred. No. 1e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMDYS 11  
I:| | | | |  
Db 445 VIPPGMDFS 453

## RESULT 4

US-10-217-700-11  
; Sequence 11, Application US/10217700  
; Publication No. US20030070191A1  
; GENERAL INFORMATION:  
; APPLICANT: Haigler, Candace H.  
; APPLICANT: Holaday, A. Scott  
; TITLE OF INVENTION: TRANSGENIC FIBER PRODUCING PLANTS WITH INCREASED  
; FILE REFERENCE: 201304/1000  
; CURRENT APPLICATION NUMBER: US/10/217,700  
; CURRENT FILING DATE: 2002-08-12  
; EARLIER APPLICATION NUMBER: 09/394,272  
; EARLIER FILING DATE: 1999-09-10  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 11  
; LENGTH: 1083  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
US-10-217-700-11

Query Match  
Best Local Similarity 66.7%; Score 36; DB 9; Length 1083;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMDYS 11  
I:| | | | |  
Db 483 VIPPGMDFS 491

## RESULT 5

US-10-217-700-9  
; Sequence 9, Application US/10217700  
; Publication No. US20030070191A1  
; GENERAL INFORMATION:  
; APPLICANT: Haigler, Candace H.  
; APPLICANT: Holaday, A. Scott  
; TITLE OF INVENTION: TRANSGENIC FIBER PRODUCING PLANTS WITH INCREASED  
; FILE REFERENCE: 201304/1000  
; CURRENT APPLICATION NUMBER: US/10/217,700  
; CURRENT FILING DATE: 2002-08-12  
; EARLIER APPLICATION NUMBER: 09/394,272  
; EARLIER FILING DATE: 1999-09-10  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 9  
; LENGTH: 1084  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
US-10-217-700-9

Query Match  
Best Local Similarity 66.7%; Score 36; DB 9; Length 1084;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMDYS 11  
I:| | | | |  
Db 453 VIPPGMDFS 461

## RESULT 6

US-09-813-408-27  
; Sequence 27, Application US/09813408

us-09-909-164-8.rapb

Wed Jun 11 15:44:48 2003

```

Publication No. US20030049619A1
GENERAL INFORMATION:
APPLICANT: Delagrave, Simon
APPLICANT: Maris, Barry
TITLE OF INVENTION: Methods For The Synthesis Of Polynucleotides And Combinatorial Libraries
TITLE OF INVENTION: Of Polynucleotides
FILE REFERENCE: HER0041
CURRENT APPLICATION NUMBER: US/09/813,408
CURRENT FILING DATE: 2001-03-21
NUMBER OF SEQ ID NOS: 85
SOFTWARE: PatentIn version 3.0
SEQ ID NO 27
LENGTH: 440
TYPE: PRT
ORGANISM: Aeropyrum pernix
US-09-813-408-27

Query Match          64.8%; Score 35; DB 9; Length 440;
Best Local Similarity 66.7%; Pred. No. 58;
Matches              6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      2 EVVPGMDY 10
DB      120 EVLPWGVY 128

RESULT 7
US-09-815-242-5111
Sequence 5111, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in Prokaryotes
TITLE OF INVENTION: ELITRA.011A
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 5111
LENGTH: 1062
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-815-242-5111

Query Match          63.0%; Score 34; DB 10; Length 1062;
Best Local Similarity 85.7%; Pred. No. 2.4e+02;
Matches              0; Mismatches 1; Indels 0; Gaps 0;

QY      5 PXGMDYS 11
DB      321 PQGMDYS 327

Publication No. US20030049619A1
GENERAL INFORMATION:
APPLICANT: Delagrave, Simon
APPLICANT: Maris, Barry
TITLE OF INVENTION: Methods For The Synthesis Of Polynucleotides And Combinatorial Libraries
TITLE OF INVENTION: Of Polynucleotides
FILE REFERENCE: HER0041
CURRENT APPLICATION NUMBER: US/09/813,408
CURRENT FILING DATE: 2001-03-21
NUMBER OF SEQ ID NOS: 85
SOFTWARE: PatentIn version 3.0
SEQ ID NO 27
LENGTH: 440
TYPE: PRT
ORGANISM: Aeropyrum pernix
US-09-813-408-27

Query Match          64.8%; Score 35; DB 9; Length 440;
Best Local Similarity 66.7%; Pred. No. 58;
Matches              6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      2 EVVPGMDY 10
DB      120 EVLPWGVY 128

RESULT 8
US-10-027-806-4
Sequence 4, Application US/10027806
Patent No. US20020160476A1
GENERAL INFORMATION:
APPLICANT: Swanson, Ronald V.
APPLICANT: Feldman, Robert A.
APPLICANT: Schleper, Christa
TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM CENARCHAEUM SYMBIOSUM
FILE REFERENCE: DCRP.002A
CURRENT APPLICATION NUMBER: US/10/027,806
CURRENT FILING DATE: 2001-12-21
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/408,020
PRIOR FILING DATE: EARLIER FILING DATE: 1999-09-29
NUMBER OF SEQ ID NOS: 123
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 4
LENGTH: 3472
TYPE: PRT
ORGANISM: Cenarchaeum symbiosum
US-10-027-806-4

Query Match          63.0%; Score 34; DB 9; Length 3472;
Best Local Similarity 45.5%; Pred. No. 9.1e+02;
Matches              5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY      1 EEVVPXGMDYS 11
DB      2294 EDVIPRGISFS 2304

RESULT 9
US-10-034-623-4
Sequence 4, Application US/10034623
Patent No. US20020198365A1
GENERAL INFORMATION:
APPLICANT: Swanson, Ronald V.
APPLICANT: Feldman, Robert A.
APPLICANT: Schleper, Christa
TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM CENARCHAEUM SYMBIOSUM
FILE REFERENCE: DCRP.002A
CURRENT APPLICATION NUMBER: US/10/034,623
CURRENT FILING DATE: 2001-12-21
PRIOR APPLICATION NUMBER: 09/408,020
PRIOR FILING DATE: 1999-09-29
PRIOR APPLICATION NUMBER: 60/102,294
PRIOR FILING DATE: 1998-09-29
NUMBER OF SEQ ID NOS: 123
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 4
LENGTH: 3472
TYPE: PRT
ORGANISM: Cenarchaeum symbiosum
US-10-034-623-4

Query Match          63.0%; Score 34; DB 9; Length 3472;
Best Local Similarity 45.5%; Pred. No. 9.1e+02;
Matches              5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY      1 EEVVPXGMDYS 11
DB      2294 EDVIPRGISFS 2304

RESULT 10
US-10-027-801-4
Sequence 4, Application US/10027801
Patent No. US20030054364A1
GENERAL INFORMATION:
APPLICANT: Swanson, Ronald V.
APPLICANT: Feldman, Robert A.
US-10-027-801-4

Query Match          63.0%; Score 34; DB 9; Length 3472;
Best Local Similarity 45.5%; Pred. No. 9.1e+02;
Matches              5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY      1 EEVVPXGMDYS 11
DB      2294 EDVIPRGISFS 2304

```

; APPLICANT: Schleper, Christa  
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM CENARCHAEUM SYMBIOSUM  
; FILE REFERENCE: DCOIP.002A  
; CURRENT APPLICATION NUMBER: US/10/027,801  
; PRIOR FILING DATE: 2001-12-21  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/408,020  
; NUMBER OF SEQ ID NOS: 123  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 4  
; LENGTH: 3472  
; TYPE: PRT  
; ORGANISM: Cenarchaeum symbiosum.  
US-10-027-801-4

Query Match  
Best Local Similarity 63.0%; Score 34; DB 9; Length 3472;  
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;  
QY 1 EEVVPXGMDYS 11  
|:|:| |:|  
Db 2294 EDVIPRGISFS 2304

RESULT 11  
US-09-948-080-14  
; Sequence 14, Application US/09948080  
; Patent No. US2002010702A1  
; GENERAL INFORMATION:  
; APPLICANT: VAN DER OSTEN, CLAUS  
; APPLICANT: HALKIER, TORDEN  
; APPLICANT: ANDERSEN, CARSTEN  
; APPLICANT: BAUDITZ, PETER  
; APPLICANT: HANSEN, PETER KAMP  
; TITLE OF INVENTION: PROTEASE VARIANTS AND COMPOSITIONS  
; FILE REFERENCE: 4946,200-US  
; CURRENT APPLICATION NUMBER: US/09/948,080  
; PRIOR FILING DATE: 2001-09-06  
; PRIOR APPLICATION NUMBER: US/08/963,851  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 14  
; LENGTH: 59  
; TYPE: PRT  
; ORGANISM: Enterococcus faecalis  
US-09-948-080-14

Query Match  
Best Local Similarity 61.1%; Score 33; DB 10; Length 59;  
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;  
QY 1 EEVVPXGMDYS 11  
|:|:| |:|  
Db 38 EKHIPGGLEYS 48

RESULT 12  
US-09-738-626-4881  
; Sequence 4881, Application US/09738626  
; Publication No. US20020197605A1  
; GENERAL INFORMATION:  
; APPLICANT: NAKAGAWA, SATOSHI  
; APPLICANT: MIZOGUCHI, HIROSHI  
; APPLICANT: ANDO, SEIKO  
; APPLICANT: HAYASHI, MIKIRO  
; APPLICANT: OCHIAI, KEIKO  
; APPLICANT: YOKOI, HARUHIKO  
; APPLICANT: TATEISHI, NAKO  
; APPLICANT: SENO, AKIHIRO  
; APPLICANT: IKEDA, MASATO  
; APPLICANT: OZAKI, AKIO  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES

; FILE REFERENCE: 249-125  
; CURRENT APPLICATION NUMBER: US/09/738,626  
; PRIOR FILING DATE: 2000-12-18  
; PRIOR APPLICATION NUMBER: JP 99/377484  
; PRIOR FILING DATE: 1999-12-16  
; PRIOR APPLICATION NUMBER: JP 00/159162  
; PRIOR FILING DATE: 2000-04-07  
; PRIOR APPLICATION NUMBER: JP 00/280988  
; NUMBER OF SEQ ID NOS: 7059  
; SOFTWARE: PatentIn ver. 3.0  
; SEQ ID NO 4881  
; LENGTH: 283  
; TYPE: PRT  
; ORGANISM: Corynebacterium glutamicum  
US-09-738-626-4881

Query Match  
Best Local Similarity 61.1%; Score 33; DB 9; Length 283;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 4 VPXGMDYS 11  
|:|:| |:|  
Db 56 VPAGADYS 63

RESULT 13  
US-09-815-242-10697  
; Sequence 10697, Application US/09815242  
; Patent No. US20020061569A1  
; GENERAL INFORMATION:  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari L.  
; APPLICANT: Zyskind, Judith W.  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John D.  
; APPLICANT: Carr, Grant J.  
; APPLICANT: Yamamoto, Robert T.  
; APPLICANT: Xu, H. Howard  
; TITLE OF INVENTION: Identification of Essential Genes In  
; FILE REFERENCE: ELITRA.011A  
; CURRENT APPLICATION NUMBER: US/09/815,242  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/269,308  
; NUMBER OF SEQ ID NOS: 14110  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 10697  
; LENGTH: 299  
; TYPE: PRT  
; ORGANISM: Enterococcus faecalis  
US-09-815-242-10697

Query Match  
Best Local Similarity 61.1%; Score 33; DB 10; Length 299;  
Matches 4; Conservative 4; Mismatches 2; Indels 0; Gaps 0;  
QY 1 EEVVPXGMDY 10  
|:|:| |:|  
Db 218 EQTPTGIEY 227

## RESULT 14

US-09-978-295A-526  
; Sequence 526, Application US/09978295A  
; Patent No. US20020156006A1  
; GENERAL INFORMATION:  
; APPLICANT: Ashkenazi, Avi  
; APPLICANT: Baker Kevin P.  
; APPLICANT: Botstein, David  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Eaton, Dan  
; APPLICANT: Ferrara, Napoleon  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Fong, Sherman  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerber, Hanspeter  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, J. Christopher  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Hillan, Kenneth J.  
; APPLICANT: Kljavin, Ivar J.  
; APPLICANT: Kuo, Sophia S.  
; APPLICANT: Napier, Mary A.  
; APPLICANT: Pan, James;  
; APPLICANT: Paoni, Nicholas F.  
; APPLICANT: Roy, Margaret Ann  
; APPLICANT: Shelton, David L.  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wood, William I.  
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
; FILE REFERENCE: P2630PIC11  
; CURRENT APPLICATION NUMBER: US/09/978, 295A  
; CURRENT FILING DATE: 2001-10-15  
; PRIOR APPLICATION NUMBER: 09/918585  
; PRIOR FILING DATE: 2001-07-30  
; PRIOR APPLICATION NUMBER: 60/062250  
; PRIOR FILING DATE: 1997-10-17  
; PRIOR APPLICATION NUMBER: 60/064249  
; PRIOR FILING DATE: 1997-11-03  
; PRIOR APPLICATION NUMBER: 60/065311  
; PRIOR FILING DATE: 1997-11-13  
; PRIOR APPLICATION NUMBER: 60/066364  
; PRIOR FILING DATE: 1997-11-21  
; PRIOR APPLICATION NUMBER: 60/077450  
; PRIOR FILING DATE: 1998-03-10  
; PRIOR APPLICATION NUMBER: 60/077632  
; PRIOR FILING DATE: 1998-03-11  
; PRIOR APPLICATION NUMBER: 60/077641  
; PRIOR FILING DATE: 1998-03-11  
; PRIOR APPLICATION NUMBER: 60/077649  
; PRIOR FILING DATE: 1998-03-11  
; PRIOR APPLICATION NUMBER: 60/077791  
; PRIOR FILING DATE: 1998-03-12  
; PRIOR APPLICATION NUMBER: 60/078004  
; PRIOR FILING DATE: 1998-03-13  
; PRIOR APPLICATION NUMBER: 60/078886  
; PRIOR FILING DATE: 1998-03-20  
; PRIOR APPLICATION NUMBER: 60/078936  
; PRIOR FILING DATE: 1998-03-20  
; PRIOR APPLICATION NUMBER: 60/078910  
; PRIOR FILING DATE: 1998-03-20  
; PRIOR APPLICATION NUMBER: 60/078939  
; PRIOR FILING DATE: 1998-03-20  
; PRIOR APPLICATION NUMBER: 60/079294  
; PRIOR FILING DATE: 1998-03-25  
; PRIOR APPLICATION NUMBER: 60/079656  
; PRIOR FILING DATE: 1998-03-26

; PRIOR APPLICATION NUMBER: 60/079664  
; PRIOR FILING DATE: 1998-03-27  
; PRIOR APPLICATION NUMBER: 60/079689  
; PRIOR FILING DATE: 1998-03-27  
; PRIOR APPLICATION NUMBER: 60/079663  
; PRIOR FILING DATE: 1998-03-27  
; PRIOR APPLICATION NUMBER: 60/079728  
; PRIOR FILING DATE: 1998-03-27  
; PRIOR APPLICATION NUMBER: 60/079786  
; PRIOR FILING DATE: 1998-03-27  
; PRIOR APPLICATION NUMBER: 60/079920  
; PRIOR FILING DATE: 1998-03-30  
; PRIOR APPLICATION NUMBER: 60/079923  
; PRIOR FILING DATE: 1998-03-30  
; PRIOR APPLICATION NUMBER: 60/080105  
; PRIOR FILING DATE: 1998-03-31  
; PRIOR APPLICATION NUMBER: 60/080107  
; PRIOR FILING DATE: 1998-03-31  
; PRIOR APPLICATION NUMBER: 60/080165  
; PRIOR FILING DATE: 1998-03-31  
; PRIOR APPLICATION NUMBER: 60/080194  
; PRIOR FILING DATE: 1998-03-31  
; PRIOR APPLICATION NUMBER: 60/080327  
; PRIOR FILING DATE: 1998-04-01  
; PRIOR APPLICATION NUMBER: 60/080328  
; PRIOR FILING DATE: 1998-04-01  
; PRIOR APPLICATION NUMBER: 60/080333  
; PRIOR FILING DATE: 1998-04-01  
; PRIOR APPLICATION NUMBER: 60/080334  
; PRIOR FILING DATE: 1998-04-01  
; PRIOR APPLICATION NUMBER: 60/081070  
; PRIOR FILING DATE: 1998-04-08  
; PRIOR APPLICATION NUMBER: 60/081049  
; PRIOR FILING DATE: 1998-04-08  
; PRIOR APPLICATION NUMBER: 60/081071  
; PRIOR FILING DATE: 1998-04-08  
; PRIOR APPLICATION NUMBER: 60/081195  
; PRIOR FILING DATE: 1998-04-08  
; PRIOR APPLICATION NUMBER: 60/081203  
; PRIOR FILING DATE: 1998-04-09  
; PRIOR APPLICATION NUMBER: 60/081229  
; PRIOR FILING DATE: 1998-04-09  
; PRIOR APPLICATION NUMBER: 60/081955  
; PRIOR FILING DATE: 1998-04-15  
; PRIOR APPLICATION NUMBER: 60/081817  
; PRIOR FILING DATE: 1998-04-15  
; PRIOR APPLICATION NUMBER: 60/081819  
; PRIOR FILING DATE: 1998-04-15  
; PRIOR APPLICATION NUMBER: 60/081952  
; PRIOR FILING DATE: 1998-04-15  
; PRIOR APPLICATION NUMBER: 60/081838  
; PRIOR FILING DATE: 1998-04-15  
; PRIOR APPLICATION NUMBER: 60/082568  
; PRIOR FILING DATE: 1998-04-21  
; PRIOR APPLICATION NUMBER: 60/082569  
; PRIOR FILING DATE: 1998-04-21  
; PRIOR APPLICATION NUMBER: 60/082704  
; PRIOR FILING DATE: 1998-04-22  
; PRIOR APPLICATION NUMBER: 60/082804  
; PRIOR FILING DATE: 1998-04-22  
; PRIOR APPLICATION NUMBER: 60/082700  
; PRIOR FILING DATE: 1998-04-22  
; PRIOR APPLICATION NUMBER: 60/082797  
; PRIOR FILING DATE: 1998-04-22  
; PRIOR APPLICATION NUMBER: 60/082796  
; PRIOR FILING DATE: 1998-04-23  
; PRIOR APPLICATION NUMBER: 60/083336  
; PRIOR FILING DATE: 1998-04-27  
; PRIOR APPLICATION NUMBER: 60/083322  
; PRIOR FILING DATE: 1998-04-28  
; PRIOR APPLICATION NUMBER: 60/083392  
; PRIOR FILING DATE: 1998-04-29  
; PRIOR APPLICATION NUMBER: 60/083495

; PRIOR FILING DATE: 1998-04-29  
; PRIOR APPLICATION NUMBER: 60/083496  
; PRIOR FILING DATE: 1998-04-29  
; PRIOR APPLICATION NUMBER: 60/083499  
; PRIOR FILING DATE: 1998-04-29  
; PRIOR APPLICATION NUMBER: 60/083545  
; PRIOR FILING DATE: 1998-04-29  
; PRIOR APPLICATION NUMBER: 60/083554  
; PRIOR FILING DATE: 1998-04-29  
; PRIOR APPLICATION NUMBER: 60/083558  
; PRIOR FILING DATE: 1998-04-29  
; PRIOR APPLICATION NUMBER: 60/083559  
; PRIOR FILING DATE: 1998-04-29  
; PRIOR APPLICATION NUMBER: 60/083500  
; PRIOR FILING DATE: 1998-04-29  
; PRIOR APPLICATION NUMBER: 60/083742  
; PRIOR FILING DATE: 1998-04-30  
; PRIOR APPLICATION NUMBER: 60/084366  
; PRIOR FILING DATE: 1998-05-05  
; PRIOR APPLICATION NUMBER: 60/084414  
; PRIOR FILING DATE: 1998-05-06  
; PRIOR APPLICATION NUMBER: 60/084441  
; PRIOR FILING DATE: 1998-05-06  
; PRIOR APPLICATION NUMBER: 60/084637  
; PRIOR FILING DATE: 1998-05-07  
; PRIOR APPLICATION NUMBER: 60/084639  
; PRIOR FILING DATE: 1998-05-07  
; PRIOR APPLICATION NUMBER: 60/084640  
; PRIOR FILING DATE: 1998-05-07  
; PRIOR APPLICATION NUMBER: 60/084598  
; PRIOR FILING DATE: 1998-05-07  
; PRIOR APPLICATION NUMBER: 60/084600  
; PRIOR FILING DATE: 1998-05-07  
; PRIOR APPLICATION NUMBER: 60/084627  
; PRIOR FILING DATE: 1998-05-07  
; PRIOR APPLICATION NUMBER: 60/084643  
; PRIOR FILING DATE: 1998-05-07  
; PRIOR APPLICATION NUMBER: 60/085339  
; PRIOR FILING DATE: 1998-05-13  
; PRIOR APPLICATION NUMBER: 60/085338  
; PRIOR FILING DATE: 1998-05-13  
; PRIOR APPLICATION NUMBER: 60/085323  
; PRIOR FILING DATE: 1998-05-13  
; PRIOR APPLICATION NUMBER: 60/085582  
; PRIOR FILING DATE: 1998-05-15  
; PRIOR APPLICATION NUMBER: 60/085700  
; PRIOR FILING DATE: 1998-05-15  
; PRIOR APPLICATION NUMBER: 60/085689  
; PRIOR FILING DATE: 1998-05-15  
; PRIOR APPLICATION NUMBER: 60/085579  
; PRIOR FILING DATE: 1998-05-15  
; PRIOR APPLICATION NUMBER: 60/085580  
; PRIOR FILING DATE: 1998-05-15  
; PRIOR APPLICATION NUMBER: 60/085573  
; PRIOR FILING DATE: 1998-05-15  
; PRIOR APPLICATION NUMBER: 60/085704  
; PRIOR FILING DATE: 1998-05-15  
; PRIOR APPLICATION NUMBER: 60/085697

Query Match 61.1%; Score 33; DB 9; Length 736;  
Best Local Similarity 70.0%; Pred. No. 2.6e+02;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMDY 10  
| | | | |  
Db 331 EPVVGMDY 340

## RESULT 15

US-09-978-697-526  
; Sequence 526, Application US/09978697  
; Patent No. US20020169284A1  
; GENERAL INFORMATION:

; APPLICANT: Ashkenazi, Avi  
; APPLICANT: Baker Kevin P.  
; APPLICANT: Botstein, David  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Eaton, Dan  
; APPLICANT: Ferrara, Napoleon  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Fong, Sherman  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerber, Hanspeter  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, J. Christopher  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Hillan, Kenneth J.  
; APPLICANT: Kljavin, Ivar J.  
; APPLICANT: Kuo, Sophia S.  
; APPLICANT: Napier, Mary A.  
; APPLICANT: Pan, James;  
; APPLICANT: Paoni, Nicholas F.  
; APPLICANT: Roy, Margaret Ann  
; APPLICANT: Shelton, David L.  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wood, William I.  
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
; FILE REFERENCE: P2630PIC27  
; CURRENT APPLICATION NUMBER: US/09/978,697  
; CURRENT FILING DATE: 2001-10-16  
; PRIOR APPLICATION NUMBER: 09/918585  
; PRIOR FILING DATE: 2001-07-30  
; PRIOR APPLICATION NUMBER: 60/062250  
; PRIOR FILING DATE: 1997-10-17  
; PRIOR APPLICATION NUMBER: 60/064249  
; PRIOR FILING DATE: 1997-11-03  
; PRIOR APPLICATION NUMBER: 60/065311  
; PRIOR FILING DATE: 1997-11-13  
; PRIOR APPLICATION NUMBER: 60/066364  
; PRIOR FILING DATE: 1997-11-21  
; PRIOR APPLICATION NUMBER: 60/077450  
; PRIOR FILING DATE: 1998-03-10  
; PRIOR APPLICATION NUMBER: 60/077632  
; PRIOR FILING DATE: 1998-03-11  
; PRIOR APPLICATION NUMBER: 60/077641  
; PRIOR FILING DATE: 1998-03-11  
; PRIOR APPLICATION NUMBER: 60/077649  
; PRIOR FILING DATE: 1998-03-11  
; PRIOR APPLICATION NUMBER: 60/077791  
; PRIOR FILING DATE: 1998-03-12  
; PRIOR APPLICATION NUMBER: 60/078004  
; PRIOR FILING DATE: 1998-03-13  
; PRIOR APPLICATION NUMBER: 60/078886  
; PRIOR FILING DATE: 1998-03-20  
; PRIOR APPLICATION NUMBER: 60/078936  
; PRIOR FILING DATE: 1998-03-20  
; PRIOR APPLICATION NUMBER: 60/078910  
; PRIOR FILING DATE: 1998-03-20  
; PRIOR APPLICATION NUMBER: 60/078939  
; PRIOR FILING DATE: 1998-03-20  
; PRIOR APPLICATION NUMBER: 60/079294  
; PRIOR FILING DATE: 1998-03-25  
; PRIOR APPLICATION NUMBER: 60/079656  
; PRIOR FILING DATE: 1998-03-26  
; PRIOR APPLICATION NUMBER: 60/079664  
; PRIOR FILING DATE: 1998-03-27  
; PRIOR APPLICATION NUMBER: 60/079689  
; PRIOR FILING DATE: 1998-03-27  
; PRIOR APPLICATION NUMBER: 60/079663  
; PRIOR FILING DATE: 1998-03-27  
; PRIOR APPLICATION NUMBER: 60/079728

**us-09-909-164-8.rapb**

PRIOR APPLICATION NUMBER: 60/083554
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083558
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083559
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083500
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083742
PRIOR FILING DATE: 1998-04-30
PRIOR APPLICATION NUMBER: 60/084366
PRIOR FILING DATE: 1998-05-05
PRIOR APPLICATION NUMBER: 60/084414
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084441
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084637
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084639
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084640
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084598
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-5-07
PRIOR APPLICATION NUMBER: 60/084627
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084643
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/085339
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085338
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085323
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085582
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085700
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085689
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085579
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085580
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085573
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085704
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085697

Query Match 61.1%; Score 33; DB 9; Length 736;  
Best Local Similarity 70.0%; Pred. No. 2.6e+02;  
Matches 7: Conservative 0; Mismatches 3; Indels

QY 1 EEVVPXGMDY 10  
| | | | |  
Db 331 EPVVVYGMDY 340

```

RESULT 16
US-09-978-192A-526
; Sequence 526, Application US/09978192A
; Patent No. US2002017753A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnovers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
;

```

APPLICANT: Fong, Sherman  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Gerber, Hanspeter  
APPLICANT: Gerritsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Grimaldi, J. Christopher  
APPLICANT: Gurney, Austin L.  
APPLICANT: Hillan, Kenneth J.  
APPLICANT: Kljavin, Ivar J.  
APPLICANT: Kuo, Sophia S.  
APPLICANT: Napier, Mary A.  
APPLICANT: Pan, James  
APPLICANT: Paoni, Nicholas F.  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Shelton, David L.  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William I.  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
TITLE OF INVENTION: Acids Encoding the Same  
FILE REFERENCE: P2630P1C9  
CURRENT APPLICATION NUMBER: US/09/978,192A  
CURRENT FILING DATE: 2001-10-15  
PRIOR APPLICATION NUMBER: 09/918585  
PRIOR FILING DATE: 2001-07-30  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/064249  
PRIOR FILING DATE: 1997-11-03  
PRIOR APPLICATION NUMBER: 60/065311  
PRIOR FILING DATE: 1997-11-13  
PRIOR APPLICATION NUMBER: 60/066364  
PRIOR FILING DATE: 1997-11-21  
PRIOR APPLICATION NUMBER: 60/077450  
PRIOR FILING DATE: 1998-03-10  
PRIOR APPLICATION NUMBER: 60/077632  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077641  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077649  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077791  
PRIOR FILING DATE: 1998-03-12  
PRIOR APPLICATION NUMBER: 60/078004  
PRIOR FILING DATE: 1998-03-13  
PRIOR APPLICATION NUMBER: 60/078886  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078936  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078939  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/079294  
PRIOR FILING DATE: 1998-03-25  
PRIOR APPLICATION NUMBER: 60/079656  
PRIOR FILING DATE: 1998-03-26  
PRIOR APPLICATION NUMBER: 60/079664  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079689  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079663  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079728  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079786  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079920  
PRIOR FILING DATE: 1998-03-30  
PRIOR APPLICATION NUMBER: 60/079923  
PRIOR FILING DATE: 1998-03-30  
PRIOR APPLICATION NUMBER: 60/080105  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080107  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080165  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080194  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080327  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080328  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080333  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080334  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/081070  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081049  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081071  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081195  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081203  
PRIOR FILING DATE: 1998-04-09  
PRIOR APPLICATION NUMBER: 60/081229  
PRIOR FILING DATE: 1998-04-09  
PRIOR APPLICATION NUMBER: 60/081955  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081817  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081819  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081952  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081838  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/082568  
PRIOR FILING DATE: 1998-04-21  
PRIOR APPLICATION NUMBER: 60/082569  
PRIOR FILING DATE: 1998-04-21  
PRIOR APPLICATION NUMBER: 60/082704  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082804  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082700  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082797  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082796  
PRIOR FILING DATE: 1998-04-23  
PRIOR APPLICATION NUMBER: 60/083336  
PRIOR FILING DATE: 1998-04-27  
PRIOR APPLICATION NUMBER: 60/083322  
PRIOR FILING DATE: 1998-04-28  
PRIOR APPLICATION NUMBER: 60/083392  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083495  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083496  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083499  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083545  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083554  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083558  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083559  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083500



GENERAL INFORMATION:

APPLICANT:	ASHKENAZI, Avi
APPLICANT:	Baker Kevin P.
APPLICANT:	Botstein, David
APPLICANT:	Desnoyers, Luc
APPLICANT:	Eaton, Dan
APPLICANT:	Ferraro, Napoleon
APPLICANT:	Filvaroff, Ellen
APPLICANT:	Fong, Sherman
APPLICANT:	Gao, Wei-Qiang
APPLICANT:	Gerber, Hanspeter
APPLICANT:	Gerritsen, Mary E.
APPLICANT:	Goddard, Audrey
APPLICANT:	Godowski, Paul J.
APPLICANT:	Grimaldi, J. Chris

; PRIOR FILING DATE: 1998-03-31  
; PRIOR APPLICATION NUMBER: 60/080327  
; PRIOR FILING DATE: 1998-04-01  
; PRIOR APPLICATION NUMBER: 60/080328  
; PRIOR FILING DATE: 1998-04-01  
; PRIOR APPLICATION NUMBER: 60/080333  
; PRIOR FILING DATE: 1998-04-01  
; PRIOR APPLICATION NUMBER: 60/080334  
; PRIOR FILING DATE: 1998-04-01  
; PRIOR APPLICATION NUMBER: 60/081070  
; PRIOR FILING DATE: 1998-04-08  
; PRIOR APPLICATION NUMBER: 60/081049  
; PRIOR FILING DATE: 1998-04-08  
; PRIOR APPLICATION NUMBER: 60/081071  
; PRIOR FILING DATE: 1998-04-08  
; PRIOR APPLICATION NUMBER: 60/081195  
; PRIOR FILING DATE: 1998-04-08  
; PRIOR APPLICATION NUMBER: 60/081203  
; PRIOR FILING DATE: 1998-04-09  
; PRIOR APPLICATION NUMBER: 60/081229  
; PRIOR FILING DATE: 1998-04-09  
; PRIOR APPLICATION NUMBER: 60/081955  
; PRIOR FILING DATE: 1998-04-15  
; PRIOR APPLICATION NUMBER: 60/081817  
; PRIOR FILING DATE: 1998-04-15  
; PRIOR APPLICATION NUMBER: 60/081819  
; PRIOR FILING DATE: 1998-04-15  
; PRIOR APPLICATION NUMBER: 60/081952  
; PRIOR FILING DATE: 1998-04-15  
; PRIOR APPLICATION NUMBER: 60/081838  
; PRIOR FILING DATE: 1998-04-15  
; PRIOR APPLICATION NUMBER: 60/082568  
; PRIOR FILING DATE: 1998-04-21  
; PRIOR APPLICATION NUMBER: 60/082569  
; PRIOR FILING DATE: 1998-04-21  
; PRIOR APPLICATION NUMBER: 60/082704  
; PRIOR FILING DATE: 1998-04-22  
; PRIOR APPLICATION NUMBER: 60/082804  
; PRIOR FILING DATE: 1998-04-22  
; PRIOR APPLICATION NUMBER: 60/082700  
; PRIOR FILING DATE: 1998-04-22  
; PRIOR APPLICATION NUMBER: 60/082797  
; PRIOR FILING DATE: 1998-04-22  
; PRIOR APPLICATION NUMBER: 60/082796  
; PRIOR FILING DATE: 1998-04-23  
; PRIOR APPLICATION NUMBER: 60/083336  
; PRIOR FILING DATE: 1998-04-27  
; PRIOR APPLICATION NUMBER: 60/083322  
; PRIOR FILING DATE: 1998-04-28  
; PRIOR APPLICATION NUMBER: 60/083392  
; PRIOR FILING DATE: 1998-04-29  
; PRIOR APPLICATION NUMBER: 60/083495  
; PRIOR FILING DATE: 1998-04-29  
; PRIOR APPLICATION NUMBER: 60/083496  
; PRIOR FILING DATE: 1998-04-29  
; PRIOR APPLICATION NUMBER: 60/083499  
; PRIOR FILING DATE: 1998-04-29  
; PRIOR APPLICATION NUMBER: 60/083545  
; PRIOR FILING DATE: 1998-04-29  
; PRIOR APPLICATION NUMBER: 60/083554  
; PRIOR FILING DATE: 1998-04-29  
; PRIOR APPLICATION NUMBER: 60/083558  
; PRIOR FILING DATE: 1998-04-29  
; PRIOR APPLICATION NUMBER: 60/083559  
; PRIOR FILING DATE: 1998-04-29  
; PRIOR APPLICATION NUMBER: 60/083500  
; PRIOR FILING DATE: 1998-04-29  
; PRIOR APPLICATION NUMBER: 60/083742  
; PRIOR FILING DATE: 1998-04-30  
; PRIOR APPLICATION NUMBER: 60/084366  
; PRIOR FILING DATE: 1998-05-05  
; PRIOR APPLICATION NUMBER: 60/084414  
; PRIOR FILING DATE: 1998-05-06

; PRIOR APPLICATION NUMBER: 60/084441  
; PRIOR FILING DATE: 1998-05-06  
; PRIOR APPLICATION NUMBER: 60/084637  
; PRIOR FILING DATE: 1998-05-07  
; PRIOR APPLICATION NUMBER: 60/084639  
; PRIOR FILING DATE: 1998-05-07  
; PRIOR APPLICATION NUMBER: 60/084640  
; PRIOR FILING DATE: 1998-05-07  
; PRIOR APPLICATION NUMBER: 60/084598  
; PRIOR FILING DATE: 1998-05-07  
; PRIOR APPLICATION NUMBER: 60/084600  
; PRIOR FILING DATE: 1998-05-07  
; PRIOR APPLICATION NUMBER: 60/084627  
; PRIOR FILING DATE: 1998-05-07  
; PRIOR APPLICATION NUMBER: 60/084643  
; PRIOR FILING DATE: 1998-05-07  
; PRIOR APPLICATION NUMBER: 60/085339  
; PRIOR FILING DATE: 1998-05-13  
; PRIOR APPLICATION NUMBER: 60/085338  
; PRIOR FILING DATE: 1998-05-13  
; PRIOR APPLICATION NUMBER: 60/085323  
; PRIOR FILING DATE: 1998-05-13  
; PRIOR APPLICATION NUMBER: 60/085582  
; PRIOR FILING DATE: 1998-05-15  
; PRIOR APPLICATION NUMBER: 60/085700  
; PRIOR FILING DATE: 1998-05-15  
; PRIOR APPLICATION NUMBER: 60/085689  
; PRIOR FILING DATE: 1998-05-15  
; PRIOR APPLICATION NUMBER: 60/085579  
; PRIOR FILING DATE: 1998-05-15  
; PRIOR APPLICATION NUMBER: 60/085580  
; PRIOR FILING DATE: 1998-05-15  
; PRIOR APPLICATION NUMBER: 60/085573  
; PRIOR FILING DATE: 1998-05-15  
; PRIOR APPLICATION NUMBER: 60/085704  
; PRIOR FILING DATE: 1998-05-15  
; PRIOR APPLICATION NUMBER: 60/085697

Query Match 61.1%; Score 33; DB 9; Length 736;  
Best Local Similarity 70.0%; Pred. No. 2.6e+02;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMDY 10  
| | | | |  
Db 331 EPVVVYGM DY 340

## RESULT 18

US-09-978-189-526  
; Sequence S26, Application US/09978189  
; Publication No. US20030004102A1  
; GENERAL INFORMATION:  
; APPLICANT: Ashkenazi, Avi  
; APPLICANT: Baker Kevin P.  
; APPLICANT: Botstein, David  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Eaton, Dan  
; APPLICANT: Ferrara, Napoleon  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Fong, Sherman  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerber, Hanspeter  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, J. Christopher  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Hillan, Kenneth J.  
; APPLICANT: Kljavin, Ivar J.  
; APPLICANT: Kuo, Sophia S.  
; APPLICANT: Napier, Mary A.  
; APPLICANT: Pan, James;  
; APPLICANT: Paoni, Nicholas F.

Wed Jun 11 15:44:48 2003

us-09-909-164-8.rapb

APPLICANT: Roy, Margaret Ann  
 APPLICANT: Shelton, David L.  
 APPLICANT: Stewart, Timothy A.  
 APPLICANT: Tumas, Daniel  
 APPLICANT: Williams, P. Mickey  
 APPLICANT: Wood, William I.  
 TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
 FILE REFERENCE: P2630PIC7  
 CURRENT APPLICATION NUMBER: US/09/978,189  
 CURRENT FILING DATE: 2001-10-15  
 PRIOR APPLICATION NUMBER: 09/918585  
 PRIOR FILING DATE: 2001-07-30  
 PRIOR APPLICATION NUMBER: 60/062250  
 PRIOR FILING DATE: 1997-10-17  
 PRIOR APPLICATION NUMBER: 60/064249  
 PRIOR FILING DATE: 1997-11-03  
 PRIOR APPLICATION NUMBER: 60/065311  
 PRIOR FILING DATE: 1997-11-13  
 PRIOR APPLICATION NUMBER: 60/066364  
 PRIOR FILING DATE: 1997-11-21  
 PRIOR APPLICATION NUMBER: 60/077450  
 PRIOR FILING DATE: 1998-03-10  
 PRIOR APPLICATION NUMBER: 60/077632  
 PRIOR FILING DATE: 1998-03-11  
 PRIOR APPLICATION NUMBER: 60/077641  
 PRIOR FILING DATE: 1998-03-11  
 PRIOR APPLICATION NUMBER: 60/077649  
 PRIOR FILING DATE: 1998-03-11  
 PRIOR APPLICATION NUMBER: 60/077791  
 PRIOR FILING DATE: 1998-03-12  
 PRIOR APPLICATION NUMBER: 60/078004  
 PRIOR FILING DATE: 1998-03-13  
 PRIOR APPLICATION NUMBER: 60/078886  
 PRIOR FILING DATE: 1998-03-20  
 PRIOR APPLICATION NUMBER: 60/078936  
 PRIOR FILING DATE: 1998-03-20  
 PRIOR APPLICATION NUMBER: 60/078910  
 PRIOR FILING DATE: 1998-03-20  
 PRIOR APPLICATION NUMBER: 60/078939  
 PRIOR FILING DATE: 1998-03-20  
 PRIOR APPLICATION NUMBER: 60/079294  
 PRIOR FILING DATE: 1998-03-25  
 PRIOR APPLICATION NUMBER: 60/079656  
 PRIOR FILING DATE: 1998-03-26  
 PRIOR APPLICATION NUMBER: 60/079664  
 PRIOR FILING DATE: 1998-03-27  
 PRIOR APPLICATION NUMBER: 60/079689  
 PRIOR FILING DATE: 1998-03-27  
 PRIOR APPLICATION NUMBER: 60/079663  
 PRIOR FILING DATE: 1998-03-27  
 PRIOR APPLICATION NUMBER: 60/079728  
 PRIOR FILING DATE: 1998-03-27  
 PRIOR APPLICATION NUMBER: 60/079786  
 PRIOR FILING DATE: 1998-03-27  
 PRIOR APPLICATION NUMBER: 60/079920  
 PRIOR FILING DATE: 1998-03-30  
 PRIOR APPLICATION NUMBER: 60/079923  
 PRIOR FILING DATE: 1998-03-30  
 PRIOR APPLICATION NUMBER: 60/080105  
 PRIOR FILING DATE: 1998-03-31  
 PRIOR APPLICATION NUMBER: 60/080107  
 PRIOR FILING DATE: 1998-03-31  
 PRIOR APPLICATION NUMBER: 60/080165  
 PRIOR FILING DATE: 1998-03-31  
 PRIOR APPLICATION NUMBER: 60/080194  
 PRIOR FILING DATE: 1998-03-31  
 PRIOR APPLICATION NUMBER: 60/080327  
 PRIOR FILING DATE: 1998-04-01  
 PRIOR APPLICATION NUMBER: 60/080328  
 PRIOR FILING DATE: 1998-04-01  
 PRIOR APPLICATION NUMBER: 60/080333  
 PRIOR FILING DATE: 1998-04-01  
 PRIOR APPLICATION NUMBER: 60/080334  
 PRIOR FILING DATE: 1998-04-01  
 PRIOR APPLICATION NUMBER: 60/081070  
 PRIOR FILING DATE: 1998-04-08  
 PRIOR APPLICATION NUMBER: 60/081049  
 PRIOR FILING DATE: 1998-04-08  
 PRIOR APPLICATION NUMBER: 60/081071  
 PRIOR FILING DATE: 1998-04-08  
 PRIOR APPLICATION NUMBER: 60/081195  
 PRIOR FILING DATE: 1998-04-08  
 PRIOR APPLICATION NUMBER: 60/081203  
 PRIOR FILING DATE: 1998-04-09  
 PRIOR APPLICATION NUMBER: 60/081229  
 PRIOR FILING DATE: 1998-04-09  
 PRIOR APPLICATION NUMBER: 60/081955  
 PRIOR FILING DATE: 1998-04-15  
 PRIOR APPLICATION NUMBER: 60/081817  
 PRIOR FILING DATE: 1998-04-15  
 PRIOR APPLICATION NUMBER: 60/081819  
 PRIOR FILING DATE: 1998-04-15  
 PRIOR APPLICATION NUMBER: 60/081952  
 PRIOR FILING DATE: 1998-04-15  
 PRIOR APPLICATION NUMBER: 60/081838  
 PRIOR FILING DATE: 1998-04-15  
 PRIOR APPLICATION NUMBER: 60/082568  
 PRIOR FILING DATE: 1998-04-21  
 PRIOR APPLICATION NUMBER: 60/082569  
 PRIOR FILING DATE: 1998-04-21  
 PRIOR APPLICATION NUMBER: 60/082704  
 PRIOR FILING DATE: 1998-04-22  
 PRIOR APPLICATION NUMBER: 60/082804  
 PRIOR FILING DATE: 1998-04-22  
 PRIOR APPLICATION NUMBER: 60/082700  
 PRIOR FILING DATE: 1998-04-22  
 PRIOR APPLICATION NUMBER: 60/082797  
 PRIOR FILING DATE: 1998-04-22  
 PRIOR APPLICATION NUMBER: 60/082796  
 PRIOR FILING DATE: 1998-04-23  
 PRIOR APPLICATION NUMBER: 60/083336  
 PRIOR FILING DATE: 1998-04-27  
 PRIOR APPLICATION NUMBER: 60/083322  
 PRIOR FILING DATE: 1998-04-28  
 PRIOR APPLICATION NUMBER: 60/083392  
 PRIOR FILING DATE: 1998-04-29  
 PRIOR APPLICATION NUMBER: 60/083495  
 PRIOR FILING DATE: 1998-04-29  
 PRIOR APPLICATION NUMBER: 60/083496  
 PRIOR FILING DATE: 1998-04-29  
 PRIOR APPLICATION NUMBER: 60/083499  
 PRIOR FILING DATE: 1998-04-29  
 PRIOR APPLICATION NUMBER: 60/083545  
 PRIOR FILING DATE: 1998-04-29  
 PRIOR APPLICATION NUMBER: 60/083554  
 PRIOR FILING DATE: 1998-04-29  
 PRIOR APPLICATION NUMBER: 60/083558  
 PRIOR FILING DATE: 1998-04-29  
 PRIOR APPLICATION NUMBER: 60/083559  
 PRIOR FILING DATE: 1998-04-29  
 PRIOR APPLICATION NUMBER: 60/083500  
 PRIOR FILING DATE: 1998-04-29  
 PRIOR APPLICATION NUMBER: 60/083742  
 PRIOR FILING DATE: 1998-04-30  
 PRIOR APPLICATION NUMBER: 60/084366  
 PRIOR FILING DATE: 1998-05-05  
 PRIOR APPLICATION NUMBER: 60/084414  
 PRIOR FILING DATE: 1998-05-06  
 PRIOR APPLICATION NUMBER: 60/084441  
 PRIOR FILING DATE: 1998-05-06  
 PRIOR APPLICATION NUMBER: 60/084637  
 PRIOR FILING DATE: 1998-05-07  
 PRIOR APPLICATION NUMBER: 60/084639  
 PRIOR FILING DATE: 1998-05-07  
 PRIOR APPLICATION NUMBER: 60/084640  
 PRIOR FILING DATE: 1998-05-07

;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/084598  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/084600  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/084627  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/084643  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/085339  
;; PRIOR FILING DATE: 1998-05-13  
;; PRIOR APPLICATION NUMBER: 60/085338  
;; PRIOR FILING DATE: 1998-05-13  
;; PRIOR APPLICATION NUMBER: 60/085323  
;; PRIOR FILING DATE: 1998-05-13  
;; PRIOR APPLICATION NUMBER: 60/085582  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085700  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085689  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085579  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085580  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085573  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085704  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085697

Query Match 61.1%; Score 33; DB 9; Length 736;  
Best Local Similarity 70.0%; Pred. No. 2.6e+02;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMDY 10  
| || ||||  
Db 331 EPVVVYGM DY 340

## RESULT 19

US-10-174-590-420  
; Sequence 420, Application US/10174590  
; Publication No. US20030008352A1

## GENERAL INFORMATION:

;; APPLICANT: Baker, Kevin P.  
;; APPLICANT: Chen, Jian  
;; APPLICANT: Desnoyers, Luc  
;; APPLICANT: Goddard, Audrey  
;; APPLICANT: Godowski, Paul J.  
;; APPLICANT: Gurney, Austin L.  
;; APPLICANT: Pan, James  
;; APPLICANT: Smith, Victoria  
;; APPLICANT: Watanabe, Colin K.  
;; APPLICANT: Wood, William I.  
;; APPLICANT: Zhang, Zemin

;; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
;; FILE REFERENCE: P3430R1C42  
;; CURRENT APPLICATION NUMBER: US/10/174,590  
;; PRIOR APPLICATION REMOVED - See File Wrapper or Palm  
;; NUMBER OF SEQ ID NOS: 612

;; SEQ ID NO 420  
;; LENGTH: 736

;; TYPE: PRT  
;; ORGANISM: Homo Sapien

US-10-174-590-420

Query Match 61.1%; Score 33; DB 9; Length 736;  
Best Local Similarity 70.0%; Pred. No. 2.6e+02;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMDY 10  
| || ||||  
Db 331 EPVVVYGM DY 340

## RESULT 20

US-10-176-758-420  
; Sequence 420, Application US/10176758  
; Publication No. US20030008353A1

## GENERAL INFORMATION:

;; APPLICANT: Baker, Kevin P.  
;; APPLICANT: Chen, Jian  
;; APPLICANT: Desnoyers, Luc  
;; APPLICANT: Goddard, Audrey  
;; APPLICANT: Godowski, Paul J.  
;; APPLICANT: Gurney, Austin L.  
;; APPLICANT: Pan, James  
;; APPLICANT: Smith, Victoria  
;; APPLICANT: Watanabe, Colin K.  
;; APPLICANT: Wood, William I.  
;; APPLICANT: Zhang, Zemin  
;; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
;; FILE REFERENCE: P3430R1C104  
;; CURRENT APPLICATION NUMBER: US/10/176,758  
;; CURRENT FILING DATE: 2002-06-21  
;; PRIOR APPLICATION REMOVED - See File Wrapper or Palm  
;; NUMBER OF SEQ ID NOS: 612  
;; SEQ ID NO 420  
;; LENGTH: 736

;; TYPE: PRT  
;; ORGANISM: Homo Sapien  
US-10-176-758-420

Query Match 61.1%; Score 33; DB 9; Length 736;  
Best Local Similarity 70.0%; Pred. No. 2.6e+02;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMDY 10  
| || ||||  
Db 331 EPVVVYGM DY 340

## RESULT 21

US-10-175-737-420  
; Sequence 420, Application US/10175737  
; Publication No. US20030013153A1

## GENERAL INFORMATION:

;; APPLICANT: Baker, Kevin P.  
;; APPLICANT: Chen, Jian  
;; APPLICANT: Desnoyers, Luc  
;; APPLICANT: Goddard, Audrey  
;; APPLICANT: Godowski, Paul J.  
;; APPLICANT: Gurney, Austin L.  
;; APPLICANT: Pan, James  
;; APPLICANT: Smith, Victoria  
;; APPLICANT: Watanabe, Colin K.  
;; APPLICANT: Wood, William I.  
;; APPLICANT: Zhang, Zemin

;; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
;; FILE REFERENCE: P3430R1C50  
;; CURRENT APPLICATION NUMBER: US/10/175,737  
;; CURRENT FILING DATE: 2002-06-19  
;; PRIOR APPLICATION REMOVED - See File Wrapper or Palm  
;; NUMBER OF SEQ ID NOS: 612

;; SEQ ID NO 420  
;; LENGTH: 736

;; TYPE: PRT  
;; ORGANISM: Homo Sapien

US-10-175-737-420

Query Match 61.1%; Score 33; DB 9; Length 736;

us-09-909-164-8.rapb

Wed Jun 11 15:44:48 2003

US-10-175-738-420  
 Query Match 61.1%; Score 33; DB 9; Length 736;  
 Best Local Similarity 70.0%; Pred. No. 2.6e+02;  
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 EEVVPXGMDY 10  
 | | | | |  
 Db 331 EPVVYGM DY 340

RESULT 24  
 US-10-175-752-420  
 ; Sequence 420, Application US/10175752  
 ; Publication No. US20030022295A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Baker, Kevin P.  
 ; APPLICANT: Chen, Jian  
 ; APPLICANT: Desnoyers, Luc  
 ; APPLICANT: Goddard, Audrey  
 ; APPLICANT: Godowski, Paul J.  
 ; APPLICANT: Gurney, Austin L.  
 ; APPLICANT: Pan, James  
 ; APPLICANT: Smith, Victoria  
 ; APPLICANT: Watanabe, Colin K.  
 ; APPLICANT: Wood, William I.  
 ; APPLICANT: Zhang, Zemin  
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
 ; FILE REFERENCE: P3430R1C50  
 ; CURRENT APPLICATION NUMBER: US/10/175,752  
 ; CURRENT FILING DATE: 2002-06-19  
 ; Prior Application removed - See File Wrapper or Palm  
 ; NUMBER OF SEQ ID NOS: 612  
 ; SEQ ID NO 420  
 ; LENGTH: 736  
 ; TYPE: PRT  
 ; ORGANISM: Homo Sapien  
 US-10-175-752-420

Query Match 61.1%; Score 33; DB 9; Length 736;  
 Best Local Similarity 70.0%; Pred. No. 2.6e+02;  
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 EEVVPXGMDY 10  
 | | | | |  
 Db 331 EPVVYGM DY 340

RESULT 25  
 US-10-176-482-420  
 ; Sequence 420, Application US/10176482  
 ; Publication No. US20030022296A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Baker, Kevin P.  
 ; APPLICANT: Chen, Jian  
 ; APPLICANT: Desnoyers, Luc  
 ; APPLICANT: Goddard, Audrey  
 ; APPLICANT: Godowski, Paul J.  
 ; APPLICANT: Gurney, Austin L.  
 ; APPLICANT: Pan, James  
 ; APPLICANT: Smith, Victoria  
 ; APPLICANT: Watanabe, Colin K.  
 ; APPLICANT: Wood, William I.  
 ; APPLICANT: Zhang, Zemin  
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
 ; FILE REFERENCE: P3430R1C70  
 ; CURRENT APPLICATION NUMBER: US/10/176,482  
 ; CURRENT FILING DATE: 2002-06-20  
 ; Prior Application removed - See File Wrapper or Palm  
 ; NUMBER OF SEQ ID NOS: 612  
 ; SEQ ID NO 420

US-10-173-706-420  
 Query Match 61.1%; Score 33; DB 9; Length 736;  
 Best Local Similarity 70.0%; Pred. No. 2.6e+02;  
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 EEVVPXGMDY 10  
 | | | | |  
 Db 331 EPVVYGM DY 340

RESULT 22  
 US-10-173-706-420  
 ; Sequence 420, Application US/10173706  
 ; Publication No. US20030022293A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Baker, Kevin P.  
 ; APPLICANT: Chen, Jian  
 ; APPLICANT: Desnoyers, Luc  
 ; APPLICANT: Goddard, Audrey  
 ; APPLICANT: Godowski, Paul J.  
 ; APPLICANT: Gurney, Austin L.  
 ; APPLICANT: Pan, James  
 ; APPLICANT: Smith, Victoria  
 ; APPLICANT: Watanabe, Colin K.  
 ; APPLICANT: Wood, William I.  
 ; APPLICANT: Zhang, Zemin  
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
 ; FILE REFERENCE: P3430R1C7  
 ; CURRENT APPLICATION NUMBER: US/10/173,706  
 ; CURRENT FILING DATE: 2002-06-17  
 ; Prior Application removed - See File Wrapper or Palm  
 ; NUMBER OF SEQ ID NOS: 612  
 ; SEQ ID NO 420  
 ; LENGTH: 736  
 ; TYPE: PRT  
 ; ORGANISM: Homo Sapien  
 US-10-173-706-420

Query Match 61.1%; Score 33; DB 9; Length 736;  
 Best Local Similarity 70.0%; Pred. No. 2.6e+02;  
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 EEVVPXGMDY 10  
 | | | | |  
 Db 331 EPVVYGM DY 340

RESULT 23  
 US-10-175-738-420  
 ; Sequence 420, Application US/10175738  
 ; Publication No. US20030022294A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Baker, Kevin P.  
 ; APPLICANT: Chen, Jian  
 ; APPLICANT: Desnoyers, Luc  
 ; APPLICANT: Goddard, Audrey  
 ; APPLICANT: Godowski, Paul J.  
 ; APPLICANT: Gurney, Austin L.  
 ; APPLICANT: Pan, James  
 ; APPLICANT: Smith, Victoria  
 ; APPLICANT: Watanabe, Colin K.  
 ; APPLICANT: Wood, William I.  
 ; APPLICANT: Zhang, Zemin  
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
 ; FILE REFERENCE: P3430R1C45  
 ; CURRENT APPLICATION NUMBER: US/10/175,738  
 ; CURRENT FILING DATE: 2002-06-19  
 ; Prior Application removed - See File Wrapper or Palm  
 ; NUMBER OF SEQ ID NOS: 612  
 ; SEQ ID NO 420  
 ; LENGTH: 736  
 ; TYPE: PRT  
 ; ORGANISM: Homo Sapien

; LENGTH: 736  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-176-482-420

Query Match 61.1%; Score 33; DB 9; Length 736;  
Best Local Similarity 70.0%; Pred. No. 2.6e+02;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMDY 10  
| | | | |  
Db 331 EPVVVYGMDY 340

Search completed: June 10, 2003, 14:35:42  
Job time : 16.0714 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run On: June 10, 2003, 13:31:15 ; Search time 11.2143 Seconds  
(without alignments)  
94.297 Million cell updates/sec

Title: US-09-909-164-8  
Perfect score: 54  
Sequence: 1 BEVVPXGMDYS 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues  
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 75 summaries

Database : PIR\_73:.\*  
1: pir1:.\*  
2: pir2:.\*  
3: pir3:.\*  
4: pir4:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	40	74.1	156	2	S54619
2	38	70.4	363	2	D69551
3	36	66.7	102	2	A42452
4	36	66.7	341	2	S72649
5	36	66.7	348	2	S72650
6	36	66.7	460	2	G96764
7	36	66.7	1049	2	JC4783
8	36	66.7	1068	1	JQ1329
9	36	66.7	1081	2	T09837
10	36	66.7	1083	2	T04062
11	36	66.7	1084	2	T04103
12	36	64.8	425	2	T24111
13	35	64.8	433	2	H87660
14	35	64.8	440	2	H72784
15	35	64.8	1150	2	T20173
16	35	64.8	1474	2	F69009
17	35	64.8	2747	2	B49132
18	34	63.0	99	2	S00210
19	34	63.0	155	2	S38255
20	34	63.0	168	2	S58208
21	34	63.0	290	2	D98182
22	34	63.0	290	2	AG3104
23	34	63.0	296	2	T72745
24	34	63.0	357	1	G69290
25	34	63.0	366	2	G69350
26	34	63.0	565	2	E86665
27	34	63.0	566	2	A70164
28	34	63.0	587	2	F81138
29	34	63.0	906	2	T48898

30	34	63.0	908	2	T48899
31	34	63.0	1062	2	F83335
32	34	63.0	1062	2	T30830
33	34	63.0	3472	2	T31308
34	33	61.1	97	2	A99427
35	33	61.1	128	2	A90471
36	33	61.1	172	2	S27021
37	33	61.1	184	2	E90335
38	33	61.1	225	2	S57810
39	33	61.1	247	2	A96001
40	33	61.1	257	2	A96546
41	33	61.1	262	2	F90298
42	33	61.1	267	2	C90307
43	33	61.1	276	2	C64417
44	33	61.1	283	2	G83055
45	33	61.1	299	2	E90487
46	33	61.1	299	2	H90352
47	33	61.1	307	2	F84330
48	33	61.1	394	2	F82491
49	33	61.1	421	1	DERTCM
50	33	61.1	670	2	S22293
51	33	61.1	797	2	S38579
52	33	61.1	800	1	TVH02F
53	33	61.1	800	2	A48991
54	33	61.1	801	2	I55363
55	33	61.1	806	2	A35963
56	33	61.1	840	2	AG0526
57	33	61.1	840	2	T39116
58	33	61.1	846	2	S57580
59	33	61.1	877	2	T40413
60	33	61.1	982	1	VCLJLK
61	33	61.1	1064	2	F86182
62	33	61.1	1401	2	G82336
63	33	61.1	2717	2	A34203
64	33	61.1	2831	2	T31419
65	33	61.1	2867	2	AG3481
66	32.5	60.2	472	1	B53236
67	32	59.3	97	2	JW0011
68	32	59.3	165	2	AG1272
69	32	59.3	165	2	AH1635
70	32	59.3	175	2	S36749
71	32	59.3	180	2	AG0504
72	32	59.3	231	1	T5ECB4
73	32	59.3	231	2	A90637
74	32	59.3	231	2	A85488
75	32	59.3	231	2	AB0515

## ALIGNMENTS

## RESULT 1

S54619  
hypothetical protein YOR013w - yeast (Saccharomyces cerevisiae)  
N:Alternate names: hypothetical protein O2612; hypothetical protein YOL303.3  
C:Species: Saccharomyces cerevisiae  
C>Date: 08-Jul-1995 #sequence\_revision 01-Sep-1995 #text\_change 19-Apr-2002  
C:Accession: S54619; S66879  
R:de Haan, M.; Maarse, A.C.; Grivell, L.A.  
submitted to the EMBL Data Library, May 1995  
A:Reference number: S54617  
A:Accession: S54619  
A:Molecule type: DNA  
A:Residues: 1-156 <DEH>  
A:Cross-references: EMBL:X87331; NID:g1041652; PIDN:CAA60762.1; PID:g829123  
R:de Haan, M.; Grivell, L.A.; Maarse, A.C.  
submitted to the Protein Sequence Database, July 1996  
A:Reference number: S66877  
A:Accession: S66879  
A:Molecule type: DNA  
A:Residues: 1-156 <DEW>  
A:Cross-references: EMBL:Z74920; NID:g1420109; PIDN:CAA99201.1; PID:g1420111; MIPS  
A:Experimental source: strain S288C

disease resistance  
RND multidrug effl  
hypothetical prote  
hypothetical 367K  
partial transposas  
hypothetical prote  
fibroblast growth  
hypothetical prote  
hypothetical prote  
conserved hypotet  
unknown protein li  
transposase ISC105  
transposase ISC105  
pantoate-beta-alan  
transposase ISC105  
hypothetical prote  
ferrisiderophore r  
acyl-CoA dehydroge  
zinc finger protei  
fibroblast growth  
fibroblast growth  
heparin-binding gr  
fibroblast growth  
protein-tyrosine k  
penicillin-binding  
probable sulfate p  
penicillin-binding  
sulfate permease -  
env polyprotein -  
hypothetical prote  
DNA-directed RNA p  
DNA-binding protei  
cyclic beta 1-2 gl  
cellobiose-phospho  
transcription fact  
plastoquinin - car  
thiol peroxidases  
thiol peroxidases  
transcription fact  
fimbrial chain (im  
L-ribulose-phospha  
L-ribulose-5-phosp  
L-ribulose-5-phosp  
L-ribulose-5-phosp

## C:Genetics:

A:Cross-references: SGD:S0005539

A:Map position: 15R

C:Superfamily: hypothetical protein YOR013w

Query Match 74.1%; Score 40; DB 2; Length 156;

Best Local Similarity 77.8%; Pred. No. 1; Indels 1; Gaps 0;

Matches 7; Conservative 1; Mismatches 1; Indels 1; Gaps 0;

QY 2 EVVPXGMDY 10

II:| |||

Db 50 EYVPLGMDY 58

## RESULT 2

D69551

conserved hypothetical protein AF2411 - Archaeoglobus fulgidus

C:Species: Archaeoglobus fulgidus

C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 22-Oct-1999

C:Accession: D69551

R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson

; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.

Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.

Nature 390, 364-370, 1997

A:Authors: Uterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.

Smith, H.O.; Woese, C.R.; Venter, J.C.

A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archae

A:Reference number: A69250; MUID:98049343; PMID:9389475

A:Accession: D69551

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-363 &lt;KLE&gt;

A:Cross-references: GB:AE001109; GB:AE000782; NID:g2689432; PIDN:AAB91255.1; PID:g265068

## Query Match

Best Local Similarity 70.4%; Score 38; DB 2; Length 363;

Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 BEVPXGMDYS 11

I:| |||

Db 120 ENIVPGIDFS 130

## RESULT 3

A42452

V1 protein - tobacco yellow dwarf virus (strain Australia)

C:Species: tobacco yellow dwarf virus

C:Date: 15-Jan-1993 #sequence\_revision 15-Jan-1993 #text\_change 08-Oct-1999

C:Accession: A42452

R:Morris, B.A.M.; Richardson, K.A.; Haley, A.; Zhan, X.; Thomas, J.E.

Virology 187, 633-642, 1992

A:Title: The nucleotide sequence of the infectious cloned DNA component of tobacco yell

A:Reference number: A42452; MUID:92188538; PMID:1546458

A:Accession: A42452

A:Molecule type: DNA

A:Residues: 1-102 &lt;MOR&gt;

A:Cross-references: GB:M81103; NID:g335283; PIDN:AAA47947.1; PID:g335284

## Query Match

Best Local Similarity 66.7%; Score 35; DB 2; Length 102;

Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 EVVPXGMDYS 11

II:| |||

Db 7 QVVPXGMDYS 16

## RESULT 4

S72649

sucrose-phosphate synthase (EC 2.4.1.14) isoform 2 - Citrus unshiu (fragment)

C:Species: Citrus unshiu

C:Date: 24-Oct-1998 #sequence\_revision 24-Oct-1998 #text\_change 21-Jul-2000

C:Accession: S72649

R:Komatsu, A.; Takanokura, Y.; Omura, M.; Akihama, T.

Mol. Gen. Genet. 252, 346-351, 1996

A:Title: Cloning and molecular analysis of cDNAs encoding three sucrose phosphate s

A:Reference number: S72648; MUID:96439842; PMID:8842155

A:Accession: S72649

A:Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-341 &lt;KOM&gt;

A:Cross-references: EMBL:AB006319; NID:g2588891; PIDN:BAA23215.1; PID:g2588892

A:Experimental source: fruit, cv. Miyagawa-Wase

A&gt;Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1997

C:Genetics:

A:Gene: SPS2

C:Function:

A:Description: catalyzes formation of sucrose-6-phosphate from UDPglucose and D-fru

A:Pathway: sucrose biosynthesis

C:Superfamily: sucrose-phosphate synthase; sucrose/sucrose-phosphate synthase homol

C:Keywords: glycosyltransferase; hexosyltransferase; sucrose biosynthesis

F:1-341/Domain: sucrose/sucrose-phosphate synthase homology (fragment) &lt;SSPS&gt;

Query Match 66.7%; Score 36; DB 2; Length 341;

Best Local Similarity 66.7%; Pred. No. 16;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMDYS 11

I:| |||

Db 228 VIPPGMDFS 236

## RESULT 5

S72650

sucrose-phosphate synthase (EC 2.4.1.14) isoform 3 - Citrus unshiu (fragment)

C:Species: Citrus unshiu

C:Date: 24-Oct-1998 #sequence\_revision 24-Oct-1998 #text\_change 21-Jul-2000

C:Accession: S72650

R:Komatsu, A.; Takanokura, Y.; Omura, M.; Akihama, T.

Mol. Gen. Genet. 252, 346-351, 1996

A:Title: Cloning and molecular analysis of cDNAs encoding three sucrose phosphate s

A:Reference number: S72648; MUID:96439842; PMID:8842155

A:Accession: S72650

A:Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-348 &lt;KOM&gt;

A:Cross-references: EMBL:AB006660; NID:g2351059; PIDN:BAA22071.1; PID:g2351060

A:Experimental source: fruit, cv. Miyagawa-Wase

A&gt;Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1997

C:Genetics:

A:Gene: SPS3

C:Function:

A:Description: catalyzes formation of sucrose-6-phosphate from UDPglucose and D-fru

A:Pathway: sucrose biosynthesis

C:Superfamily: sucrose-phosphate synthase; sucrose/sucrose-phosphate synthase homol

C:Keywords: glycosyltransferase; hexosyltransferase; sucrose biosynthesis

F:1-348/Domain: sucrose/sucrose-phosphate synthase homology (fragment) &lt;SSPS&gt;

Query Match 66.7%; Score 36; DB 2; Length 348;

Best Local Similarity 66.7%; Pred. No. 17;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMDYS 11

I:| |||

Db 234 VIPPGMDFS 242

## RESULT 6

G96764

unknown protein F25P22.17 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001

C:Accession: G96764

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,

ansen, N.F.; Hughes, B.; Huizar, L.



Nature 408, 816-820, 2000  
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A:Reference number: A86141; MUID:21016719; PMID:11130712  
A:Accession: G96764  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-460 <STO>  
A:Cross-references: GB:AE005173; NID:g6692750; PIDN:AAF24856.1; GSPDB:GN00141  
C:Genetics:  
A:Gene: F25P22.17  
A:Map position: 1

Query Match 66.7%; Score 36; DB 2; Length 460;  
Best Local Similarity 70.0%; Pred. No. 23;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMDY 10  
||| |||  
DB 218 EEDVPSAMDY 227

RESULT 7  
JC4783  
sucrose-phosphate synthase (EC 2.4.1.14) - rice  
C:Species: Oryza sativa (rice)  
C:Date: 10-May-1996 #sequence\_revision 16-Aug-1996 #text\_change 18-Jun-1999  
C:Accession: JC4783  
R:Valdez-Alarcon, J.J.; Ferrando, M.; Salerno, G.; Jimenez-Moralla, B.; Herrera-Estrella  
Gene 170, 217-222, 1996  
A:Title: Characterization of a rice sucrose-phosphate synthase-encoding gene.  
A:Reference number: JC4783; MUID:96235138; PMID:8666248  
A:Accession: JC4783  
A:Molecule type: mRNA  
A:Residues: 1-1049 <VAL>  
A:Cross-references: GB:U33175; NID:gl449931; PIDN:AAC49379.1; PID:g988270  
A:Note: UDPglucose-fructose-phosphate glucosyltransferase; Sucrosephosphate-UDPglucosyl  
C:Comment: This enzyme catalyzes the formation of sucrose-phosphate form UDP-glucose and  
C:Genetics:  
A:Gene: Sps1  
A:Introns: 24/1; 103/3; 183/3; 205/3; 435/3; 475/3; 519/3; 578/3; 596/3; 617/3; 931/3; 9  
C:Function:  
A:Description: catalyzes the formation of sucrose-6-phosphate from UDPglucose and D-fruc  
A:Pathway: sucrose biosynthesis  
C:Superfamily: sucrose-phosphate synthase; sucrose/sucrose-phosphate synthase homology  
C:Keywords: glycosyltransferase; hexosyltransferase; sucrose biosynthesis  
F:180-663/Domain: sucrose/sucrose-phosphate synthase homology <SSPS>

Query Match 66.7%; Score 36; DB 2; Length 1049;  
Best Local Similarity 66.7%; Pred. No. 58;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMDYS 11  
||| |||:  
DB 436 VIPPGMDFS 444

RESULT 8  
JQ1329  
sucrose-phosphate synthase (EC 2.4.1.14) - maize  
C:Species: Zea mays (maize)  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
C:Accession: JQ1329; PQ0260  
R:Worrell, A.C.; Bruneau, J.M.; Summerfelt, K.; Boersig, M.; Voelker, T.A.  
Plant Cell 3, 1121-1130, 1991  
A:Title: Expression of a maize sucrose phosphate synthase in tomato alters leaf carbohy  
A:Reference number: JQ1329; MUID:92338837; PMID:1840396  
A:Accession: JQ1329  
A:Molecule type: mRNA

A:Residues: 1-1068 <WOR>  
A:Cross-references: GB:M97550; NID:gl68625; PIDN:AAA33513.1; PID:gl68626  
A:Accession: PQ0260  
A:Molecule type: protein  
A:Residues: 71-74; 206-212; 471-481; 872-892 <WOR>  
C:Comment: This enzyme transfers the glucosyl group from UDPglucose to fructose-6-ph  
C:Comment: This enzyme is involved in the regulation of carbon partitioning in the  
C:Function:  
A:Description: catalyzes the formation of sucrose-6-phosphate from UDPglucose and D  
A:Pathway: sucrose biosynthesis  
C:Superfamily: sucrose-phosphate synthase; sucrose/sucrose-phosphate synthase homol  
C:Keywords: glycosyltransferase; hexosyltransferase; sucrose biosynthesis  
F:178-666/Domain: sucrose/sucrose-phosphate synthase homology <SSPS>

Query Match 66.7%; Score 36; DB 1; Length 1068;  
Best Local Similarity 66.7%; Pred. No. 59;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMDYS 11  
||| |||:  
DB 435 VIPPGMDFS 443

RESULT 9  
T09837  
sucrose-phosphate synthase (EC 2.4.1.14) isoform 2 - Craterostigma plantagineum  
C:Species: Craterostigma plantagineum  
C:Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 21-Jul-2000  
C:Accession: T09837  
R:Ingram, J.; Chandler, J.W.; Gallagher, L.; Salamini, F.; Bartels, D.  
Plant Physiol. 115, 113-121, 1997  
A:Title: Analysis of cDNA clones encoding sucrose-phosphate synthase in relation to  
A:Reference number: 216874; MUID:97451773; PMID:9306694  
A:Accession: T09837  
A:Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: mRNA  
A:Residues: 1-1081 <ING>  
A:Cross-references: EMBL:Y11795; NID:g2190349; PIDN:CAA72491.1; PID:g2190350  
A:Experimental source: ABA-treated callus  
C:Genetics:  
A:Gene: sps2  
C:Function:  
A:Description: catalyzes the formation of sucrose-6-phosphate from UDPglucose and  
A:Pathway: sucrose biosynthesis  
C:Superfamily: sucrose-phosphate synthase; sucrose/sucrose-phosphate synthase homol  
C:Keywords: glycosyltransferase; hexosyltransferase; sucrose biosynthesis  
F:176-674/Domain: sucrose/sucrose-phosphate synthase homology <SSPS>

Query Match 66.7%; Score 36; DB 2; Length 1081;  
Best Local Similarity 66.7%; Pred. No. 60;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMDYS 11  
||| |||:  
DB 445 VIPPGMDFS 453

RESULT 10  
T04062  
sucrose-phosphate synthase homolog F28M11.40 - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 30-Apr-1999 #sequence\_revision 30-Apr-1999 #text\_change 16-Jul-1999  
C:Accession: T04062  
R:Bevan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Mayer,  
submitted to the Protein Sequence Database, March 1999  
A:Reference number: Z15184  
A:Accession: T04062  
A:Molecule type: DNA  
A:Residues: 1-1083 <BEV>  
A:Cross-references: EMBL:AL049487  
A:Experimental source: cultivar Columbia; BAC clone F28M11  
C:Genetics:  
A:Map position: 4

A; Introns: 86/3; 116/3; 255/3; 322/2; 482/3; 526/3; 570/3; 629/3; 647/3; 668/3; 949/3; 9  
 A; Note: F28M11.40  
 C; Superfamily: sucrose-phosphate synthase; sucrose/sucrose-phosphate synthase homology  
 F: 230-714/Domain: sucrose-phosphate synthase homology <SSPS>

Query Match 66.7%; Score 36; DB 2; Length 1083;  
 Best Local Similarity 66.7%; Pred. No. 60;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMDYS 11  
 |:|:|:|:|  
 Db 483 VIPPGMDFS 491

## RESULT 11

T04103  
 sucrose-phosphate synthase (EC 2.4.1.14) 1 - rice  
 C; Species: Oryza sativa (rice)  
 C; Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 20-Jun-2000  
 C; Accession: T04103  
 R; Sakamoto, M.; Satozawa, T.; Kishimoto, N.; Higo, K.; Shimada, H.; Fujimura, T.  
 Plant Sci. 112, 207-217, 1995  
 A; Title: Structure and RFLP mapping of a rice sucrose phosphate synthase (SPS) gene that  
 A; Reference number: Z15212  
 A; Accession: T04103  
 A; Status: preliminary; translated from GB/EMBL/DBJ  
 A; Molecule type: DNA  
 A; Residues: 1-1084 <SAK>  
 A; Cross-references: EMBL:D45890; PIDN:BAA08304.1  
 A; Experimental source: subsp. Japonica  
 C; Genetics:  
 A; Gene: Sps1  
 A; Map position: 1

A; Introns: 120/3; 200/2; 221/3; 452/3; 492/3; 536/3; 595/3; 613/3; 634/3; 946/3; 989/2  
 C; Superfamily: sucrose-phosphate synthase; sucrose/sucrose-phosphate synthase homology  
 C; Keywords: glycosyltransferase; hexosyltransferase  
 F: 196-680/Domain: sucrose-phosphate synthase homology <SSPS>

Query Match 66.7%; Score 36; DB 2; Length 1084;  
 Best Local Similarity 66.7%; Pred. No. 60;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMDYS 11  
 |:|:|:|:|  
 Db 453 VIPPGMDFS 461

## RESULT 12

T24111  
 hypothetical protein R10D12.10 - Caenorhabditis elegans  
 C; Species: Caenorhabditis elegans  
 C; Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
 C; Accession: T24111  
 R; Percy, C.  
 submitted to the EMBL Data Library, October 1996  
 A; Reference number: Z19842  
 A; Accession: T24111  
 A; Status: preliminary; translated from GB/EMBL/DBJ  
 A; Molecule type: DNA  
 A; Residues: 1-425 <WIL>  
 A; Cross-references: EMBL:Z81109; PIDN:CA803241.1; GSPDB:GN00023; CESP:R10D12.10  
 A; Experimental source: clone R10D12  
 C; Genetics:  
 A; Gene: CESP:R10D12.10  
 A; Map position: 5  
 A; Introns: 23/3; 56/3; 113/3; 257/2

Query Match 64.8%; Score 35; DB 2; Length 425;  
 Best Local Similarity 50.0%; Pred. No. 34;  
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMDY 10  
 |:|:|:|:|

Db 335 EQIVPGGLQY 344

## RESULT 13

H87660  
 peptidoglycan-binding protein, probable [imported] - Caulobacter crescentus  
 C; Species: Caulobacter crescentus  
 C; Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 20-Apr-2001  
 C; Accession: H87660  
 R; Niernman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg  
 B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.;  
 n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser,  
 Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
 A; Title: Complete Genome Sequence of Caulobacter crescentus.  
 A; Reference number: A87249; MUID:21173698; PMID:11259647  
 A; Accession: H87660  
 A; Status: preliminary  
 A; Molecule type: DNA  
 A; Residues: 1-433 <STO>  
 A; Cross-references: GB:AE005673; NID:g13425020; PIDN:AAK25284.1; GSPDB:GN00148  
 C; Genetics:  
 A; Gene: CC3322

Query Match 64.8%; Score 35; DB 2; Length 433;  
 Best Local Similarity 54.5%; Pred. No. 34;  
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMDYS 11  
 |:|:|:|:|  
 Db 266 EVILPPGFDYS 276

## RESULT 14

H72784  
 probable alkaline proteinase APE0263 - Aeropyrum pernix (strain KI)  
 C; Species: Aeropyrum pernix  
 C; Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 20-Jun-2000  
 C; Accession: H72784  
 R; Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, X.;  
 awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, T.  
 DNA Res. 6, 83-101, 1999  
 A; Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aer  
 A; Reference number: A72450; MUID:99310339; PMID:10382966  
 A; Accession: H72784  
 A; Status: preliminary  
 A; Molecule type: DNA  
 A; Residues: 1-440 <KAW>  
 A; Cross-references: DDBJ:AP000058; NID:g5103388; PIDN:BAA79178.1; PID:g5103657  
 A; Experimental source: strain KI  
 C; Genetics:  
 A; Gene: APE0263  
 C; Superfamily: subtilisin; subtilisin homology

Query Match 64.8%; Score 35; DB 2; Length 440;  
 Best Local Similarity 66.7%; Pred. No. 35;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 EVVPXGMDY 10  
 |:|:|:|:|  
 Db 120 EVLPWGVGY 128

## RESULT 15

T20173  
 hypothetical protein C53A5.2 - Caenorhabditis elegans  
 C; Species: Caenorhabditis elegans  
 C; Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 29-Oct-1999  
 C; Accession: T20173; T23857  
 R; Mortimore, B.  
 submitted to the EMBL Data Library, November 1996  
 A; Reference number: Z19232  
 A; Accession: T20173  
 A; Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA  
 A:Residues: 1-1150 <WIL>  
 A:Cross-references: EMBL:Z81486; PIDN:CAB03994.1; GSPDB:GN00023; CESP:C53A5.2  
 A:Experimental source: clone C53A5  
 R:Matthews, L.  
 Submitted to the EMBL Data Library, August 1996  
 A:Reference number: Z19808  
 A:Accession: T23857  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-1150 <W12>  
 A:Cross-references: EMBL:Z78015; PIDN:CAB01437.1; GSPDB:GN00023; CESP:C53A5.2  
 A:Experimental source: clone R02D5  
 C:Genetics:  
 A:Gene: CESP:C53A5.2  
 A:Map position: 5  
 A:Introns: 33/3; 63/3; 132/3; 169/3; 221/3; 299/3; 379/2; 423/2; 438/2; 471/1; 513/2; 65

Query Match 64.8%; Score 35; DB 2; Length 1150;  
 Best Local Similarity 66.7%; Pred. No. 1e+02;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMDYS 11  
 Db 562 VLPVGIDYS 570

RESULT 16  
 F69009  
 probable membrane protein MTH1074 - Methanobacterium thermoautotrophicum (strain Delta H  
 C:Species: Methanobacterium thermoautotrophicum  
 C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 22-Oct-1999  
 C:Accession: F69009  
 R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;  
 ; Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiawani, N.  
 ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.  
 J. Bacteriol. 179, 7135-7155, 1997  
 A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funct  
 A:Reference number: A69000; MUID:98037514; PMID:9371463  
 A:Accession: F69009  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-1474 <MTH>  
 A:Cross-references: GB:AE000878; GB:AE000666; NID:g2622171; PIDN:AAB85563.1; PID:g262217  
 A:Experimental source: strain Delta H  
 C:Genetics:  
 A:Gene: MTH1074  
 C:Keywords: duplication

Query Match 64.8%; Score 35; DB 2; Length 1474;  
 Best Local Similarity 55.6%; Pred. No. 1.4e+02;  
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 EVVPXGMDY 10  
 Db 1238 DVLPAGLDY 1246

RESULT 17  
 B49132  
 fat facets (faf) splice form 1 - fruit fly (Drosophila melanogaster)  
 C:Species: Drosophila melanogaster  
 C:Date: 19-Dec-1993 #sequence\_revision 25-Apr-1997 #text\_change 01-Dec-2000  
 C:Accession: B49132; A49132  
 R:Fischer-Vize, J.A.; Rubin, G.M.; Lehmann, R.  
 Development 116, 985-1000, 1992  
 A:Title: The fat facets gene is required for Drosophila eye and embryo development.  
 A:Reference number: A49132; MUID:93202020; PMID:1295747  
 A:Contents: isogenic st  
 A:Accession: B49132  
 A:Status: preliminary  
 A:Molecule type: nucleic acid  
 A:Residues: 1-2747 <FIS>

A:Cross-references: GB:L04959; NID:g157411; PIDN:AAF01345.1; PID:g6013474  
 A:Note: sequence inconsistent with the nucleotide translation  
 A:Note: sequence extracted from NCBI backbone (NCBIN:127836, NCBIN:129008, NCBIP:12  
 A:Accession: A49132  
 A:Status: preliminary  
 A:Molecule type: nucleic acid  
 A:Residues: 1-2704, 'VT', 2707, 'ANNV' <FI2>  
 A:Cross-references: GB:L04958; NID:g157410; PIDN:AAF01346.1; PID:g6013475  
 A:Note: sequence extracted from NCBI backbone (NCBIN:127836, NCBIN:129008, NCBIP:12  
 C:Keywords: alternative splicing

Query Match 64.8%; Score 35; DB 2; Length 2747;  
 Best Local Similarity 54.5%; Pred. No. 2.7e+02;  
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMDYS 11  
 Db 1394 EVIVPGQDFS 1404

## RESULT 18

S00210  
 plastocyanin b - Lombardy poplar  
 C:Species: Populus nigra var. italica (Lombardy poplar).  
 C:Date: 31-Dec-1988 #sequence\_revision 31-Dec-1988 #text\_change 20-Apr-2000  
 C:Accession: S00210  
 R:Dimitrov, M.I.; Egorov, C.A.; Donchev, A.A.; Atanasov, B.P.  
 FEBS Lett. 226, 17-22, 1987  
 A:Title: Complete amino acid sequence of poplar plastocyanin b.  
 A:Reference number: S00210  
 A:Accession: S00210  
 A:Molecule type: protein  
 A:Residues: 1-99 <DIM>

C:Superfamily: plastocyanin  
 C:Keywords: chloroplast; copper; electron transfer; metalloprotein  
 F:37,84,87,92/Binding site: copper (His, Cys, His, Met) (type 1) #status predicted

Query Match 63.0%; Score 34; DB 2; Length 99;  
 Best Local Similarity 54.5%; Pred. NO. 11;  
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMDYS 11  
 Db 43 EDVPSGVDS 53

## RESULT 19

S38255  
 plastocyanin precursor - barley  
 C:Species: Hordeum vulgare (barley)  
 C:Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 11-Jun-1999  
 C:Accession: S38255; S00206  
 R:Nielsen, P.S.; Gausing, K.  
 Eur. J. Biochem. 217, 97-104, 1993  
 A:Title: In vitro binding of nuclear proteins to the barley plastocyanin gene prom  
 A:Reference number: S38255; MUID:94039081; PMID:8223592  
 A:Accession: S38255  
 A:Molecule type: DNA  
 A:Residues: 1-155 <NIEL>

A:Cross-references: EMBL:Z28347; NID:g431919; PIDN:CAA82201.1; PID:g431920  
 A:Experimental source: strain NK 1558  
 R:Nielsen, P.S.; Gausing, K.  
 FEBS Lett. 225, 159-162, 1987  
 A:Title: The precursor of barley plastocyanin: sequence of cDNA clones and gene ex  
 A:Reference number: S00206  
 A:Accession: S00206

A:Status: not compared with conceptual translation  
 A:Molecule type: mRNA  
 A:Residues: 1-119, 'T', 121-155 <NIE2>  
 A:Cross-references: EMBL:Y00704; NID:g22704  
 A:Note: not compared to nucleotide translation  
 C:Genetics:  
 A:Genome: nuclear

C:Function:  
A:Description: mediates the transfer of electrons from cytochrome b6/f to photosystem I  
C:Superfamily: plastocyanin  
C:Keywords: chloroplast; copper; electron transfer; metalloprotein  
F:1-56/Domain: transit peptide (chloroplast) #status: predicted <NTP>  
F:59-155/Product: plastocyanin #status: predicted <MAT>  
F:95,140,143,148/Binding site: copper (His, Cys, His, Met) (type 1) #status: predicted

Query Match 63.0%; Score 34; DB 2; Length 155;  
Best Local Similarity 54.5%; Pred. No. 17;  
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMDYS 11  
|:|:|:|:|  
Db 101 EDVPSGVDVS 111

RESULT 20  
S58208  
C:Species: Populus nigra (black poplar)  
C:Date: 13-Jan-1996 #sequence\_revision 01-Mar-1996 #text\_change 05-May-2000  
C:Accession: S58208  
R:Reichert, J.; Jenzelski, V.; Hachnel, W.  
A:Description: Kinetic studies of recombinant poplar plastocyanins.  
A:Reference number: S58208  
A:Accession: S58208  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-168 <RET>  
A:Cross-references: EMBL:Z50186; NID:G929814; PIDN:CAA90565.1; PID:G929815  
C:Superfamily: plastocyanin  
C:Keywords: copper; electron transfer; metalloprotein  
F:106,153,156,161/Binding site: copper (His, Cys, His, Met) (type 1) #status: predicted

Query Match 63.0%; Score 34; DB 2; Length 168;  
Best Local Similarity 54.5%; Pred. No. 19;  
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMDYS 11  
|:|:|:|:|  
Db 112 EDVPSGVDVS 122

RESULT 21  
D98182  
O6-methylguanine-DNA methyltransferase PA2118 [imported] - Agrobacterium tumefaciens (st  
C:Species: Agrobacterium tumefaciens  
C:Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 11-Jan-2002  
C:Accession: D98182  
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,  
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;  
Science 294, 2323-2328, 2001  
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum  
A:Reference number: A97359; PMID:11743194  
A:Accession: D98182  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-290 <KUR>  
A:Cross-references: GB:AE007870; PIDN:AAK8982.1; PID:G15158766; GSPDB:GN00170  
C:Genetics:  
A:Gene: AGR\_L\_818  
A:Map position: linear chromosome

Query Match 63.0%; Score 34; DB 2; Length 290;  
Best Local Similarity 50.0%; Pred. No. 35;  
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMDY 10  
|:|:|:|:|  
Db 9 EDITPIGSDY 18

RESULT 22  
AG3104  
6-O-methylguanine-DNA methyltransferase [imported] - Agrobacterium tumefaciens (str  
C:Species: Agrobacterium tumefaciens  
C:Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 11-Jan-2002  
C:Accession: AG3104  
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; W  
erage, G.; Gilliet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; J  
; Karp, P.; Romero, P.; Zhang, S.  
Science 294, 2317-2323, 2001  
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-K  
ster, E.W.  
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.  
A:Reference number: AB2577; PMID:11743193  
A:Accession: AG3104  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-290 <KUR>  
A:Cross-references: GB:AE008689; PIDN:AAI45253.1; PID:G17742937; GSPDB:GN00187  
A:Experimental source: strain C58 (Dupont)  
C:Genetics:  
A:Gene: ada  
A:Map position: linear chromosome

Query Match 63.0%; Score 34; DB 2; Length 290;  
Best Local Similarity 50.0%; Pred. No. 35;  
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMDY 10  
|:|:|:|:|  
Db 9 EDITPIGSDY 18

RESULT 23  
F72745  
Hypothetical protein APE0493 - Aeropyrum pernix (strain K1)  
C:Species: Aeropyrum pernix  
C:Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 09-Jun-2000  
C:Accession: F72745  
R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; T  
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudo, Y.; Yamazaki, J  
DNA Res. 6, 83-101, 1999  
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aer  
A:Reference number: A72450; MUID:99310339; PMID:10382966  
A:Accession: F72745  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-296 <KAW>  
A:Cross-references: DDBJ:AP000059; NID:G5103911; PIDN:BAA79458.1; PID:dl043244; PID:  
A:Experimental source: strain K1  
C:Genetics:  
A:Gene: APE0493  
C:Superfamily: Aeropyrum pernix hypothetical protein APE0493

Query Match 63.0%; Score 34; DB 2; Length 296;  
Best Local Similarity 50.0%; Pred. No. 36;  
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 EEVVPXGMDYS 11  
|:|:|:|:|  
Db 3 ETLPGGLDYT 12

RESULT 24  
G69290  
probable hexosyltransferase (EC 2.4.1.-) AF0327 [similarity] - Archaeoglobus fulgidus  
C:Species: Archaeoglobus fulgidus  
C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 20-Apr-2000  
C:Accession: G69290  
R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Do  
.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, J  
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.

Nature 390, 364-370, 1997  
 A:Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.  
 Smith, H.O.; Woese, C.R.; Venter, J.C.  
 A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeon  
 A:Reference number: A69250; MUID:98049343; PMID:9389475

A:Accession: G69290  
 A:Status: nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-357 <KLE>  
 A:Cross-references: GB:AE001082; GB:AE000782; NID:g2689405; PIDN:AAB90909.1; PID:g265031  
 C:Superfamily: probable hexosyltransferase ytnX  
 C:Keywords: glycosyltransferase; hexosyltransferase

Query Match 63.0%; Score 34; DB 1; Length 357;  
 Best Local Similarity 55.6%; Pred. No. 44;  
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 EVVPXGMDY 10  
 DB 170 EVIPNGIDF 178

RESULT 25

G69350  
 L-lactate dehydrogenase, cytochrome-type (lldD) homolog - Archaeoglobus fulgidus  
 C:Species: Archaeoglobus fulgidus  
 C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 17-Mar-2000  
 C:Accession: G69350  
 R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson  
 ; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.  
 Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.  
 Nature 390, 364-370, 1997  
 A:Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.  
 Smith, H.O.; Woese, C.R.; Venter, J.C.  
 A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeon  
 A:Reference number: A69250; MUID:98049343; PMID:9389475  
 A:Accession: G69350  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-366 <KLE>  
 A:Cross-references: GB:AE001049; GB:AE000782; NID:g2689372; PIDN:AAB90435.1; PID:g264980  
 C:Superfamily: (S)-2-hydroxy-acid oxidase; (S)-3-hydroxy-acid oxidase homology

Query Match 63.0%; Score 34; DB 2; Length 366;  
 Best Local Similarity 66.7%; Pred. No. 46;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXGMD 9  
 DB 289 EKVVTGVD 297

Search completed: June 10, 2003, 13:49:13  
 Job time : 13.2143 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2003, 13:25:04 ; Search time 4.5 Seconds  
(without alignments)  
101.387 Million cell updates/sec

Title: US-09-909-164-8  
Perfect score: 54  
Sequence: 1 BEVVPXGMDYS 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 75 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	38	70.4	1058	1 CARB_FUSNN	Q8rg86 fusobacteri
2	36	66.7	102	1 Y1LK_TYDVA	P31619 tobacco yel
3	36	66.7	1049	1 SPS_ORISA	Q43802 oryza sativ
4	36	66.7	1068	1 SPS_MAIZE	P31927 zea mays (m
5	36	66.7	1081	1 SPS2_CRAPL	O04933 craterostig
6	35	64.8	2747	1 FAF_DRONE	P55824 drosophila
7	34.5	63.9	748	1 KHL1_HUMAN	Q9nr64 homo sapien
8	34	63.0	154	1 PLAS_ORISA	P20423 oryza sativ
9	34	63.0	155	1 PLAS_HORVU	P08248 hordeum vul
10	34	63.0	168	1 PLAT_POINI	P11970 populus nig
11	34	63.0	566	1 SYFB_BORBU	P94283 borrelia bu
12	33	61.1	276	1 Y939_METJA	Q38349 methanococc
13	33	61.1	283	1 PANC_PSEAE	Q9hv69 pseudomonas
14	33	61.1	394	1 HMPA_VIBCH	Q9kmy3 vibrio chol
15	33	61.1	421	1 ACDM_RAT	P08503 rattus norv
16	33	61.1	421	1 ECB2_HALEL	O52250 halomonas e
17	33	61.1	423	1 ECB1_HALEL	Q9zeu7 halomonas e
18	33	61.1	801	1 FGR3_MOUSE	O60344 homo sapien
19	33	61.1	806	1 CEK2_CHICK	Q61851 mus musculu
20	33	61.1	806	1 CEK2_CHICK	P18460 gallus gall
21	33	61.1	877	1 SULH_SCHPO	O74377 schizosacch
22	33	61.1	982	1 ENV_SFV3L	P27399 simian foam
23	33	61.1	1401	1 ROC_VIBCH	Q9kv29 vibrio chol
24	33	61.1	2717	1 ZEP1_HUMAN	P15822 homo sapien
25	32.5	60.2	472	1 ET2A_XENLA	P19102 xenopus lae
26	32	59.3	97	1 PLAS_DAUCA	P20422 daucus caro
27	32	59.3	175	1 HES3_RAT	Q04667 rattus norv
28	32	59.3	231	1 ARAD_ECOLI	P08203 escherichia
29	32	59.3	231	1 ARAD_SALTY	P06190 salmonella
30	32	59.3	233	1 HIS9_THEMA	Q9wzr1 thermotoga
31	32	59.3	288	1 CGD2_RAT	Q04827 rattus norv
32	32	59.3	289	1 CGD2_HUMAN	P30279 mus sapien
33	32	59.3	289	1 CGD2_MOUSE	P30280 mus musculu

34	32	59.3	291	1 CGD1_BRARE	Q90459 brachydanio
35	32	59.3	291	1 CGD1_XENLA	P50755 xenopus lae
36	32	59.3	291	1 CGD2_CHICK	P49706 gallus gall
37	32	59.3	291	1 CGD2_XENLA	P53782 xenopus lae
38	32	59.3	292	1 CGD1_CHICK	P55169 gallus gall
39	32	59.3	292	1 CGD3_HUMAN	P20281 homo sapien
40	32	59.3	295	1 CGD1_HUMAN	P24385 homo sapien
41	32	59.3	295	1 CGD1_MOUSE	P25322 mus musculu
42	32	59.3	295	1 CGD1_RAT	P39948 rattus norv
43	32	59.3	338	1 MTBA_METBA	Q30640 methanosarc
44	32	59.3	472	1 ET2B_XENLA	Q91712 xenopus lae
45	32	59.3	561	1 HNFB_XENLA	Q91910 xenopus lae
46	32	59.3	561	1 SCTL_YEAST	P32784 saccharomyc
47	32	59.3	866	1 RECE_ECOLI	P15032 escherichia
48	32	59.3	995	1 HIPI_HUMAN	O00291 homo sapien
49	32	59.3	1176	1 NIR_NEUCR	P38681 neurospora
50	32	59.3	1258	1 ACN1_MOUSE	Q61137 mus musculu
51	32	59.3	1394	1 LTBS_HUMAN	P22084 homo sapien
52	32	59.3	1498	1 Y1A9_CLOAB	Q04351 clostridium
53	32	59.3	1595	1 LTBL_HUMAN	Q14766 homo sapien
54	32	59.3	1712	1 LTBL_RAT	Q09188 rattus norv
55	32	59.3	3174	1 CHAC_HUMAN	Q09188 rattus norv
56	31	57.4	98	1 PLAS_ENTPR	P07465 enteromorph
57	31	57.4	98	1 PLAS_ULVPE	P31333 ulva arsaak
58	31	57.4	98	1 PLAS_RUMOB	P56274 ulva pertus
59	31	57.4	99	1 PLAS_TOBAC	P00298 rumex obtus
60	31	57.4	99	1 REV_SIVCZ	P35476 nicotiana t
61	31	57.4	124	1 HES3_MOUSE	P17280 chimpanzee
62	31	57.4	175	1 YC10_METJA	Q61657 mus musculu
63	31	57.4	258	1 YHAI_CRYPA	Q58607 methanococc
64	31	57.4	319	1 YK14_CABEL	P10941 cryptonectr
65	31	57.4	327	1 MTBA_METAC	P34338 caenorhabdi
66	31	57.4	338	1 HYPE_BRAJA	P58869 methanosarc
67	31	57.4	346	1 LE11_METH	P31906 bradyrhizob
68	31	57.4	391	1 TOLB_HAEIN	Q27667 methanobact
69	31	57.4	427	1 LET1_KLUJA	P44677 haemophilus
70	31	57.4	469	1 ENP3_HUMAN	P53998 kluyveromyc
71	31	57.4	529	1 GUAA_MYCLE	O75355 homo sapien
72	31	57.4	529	1 GIDA_STAY3	P46810 mycobacteri
73	31	57.4	625	1 DNK1_NEUCR	Q99qt4 staphylococ
74	31	57.4	692	1 CY14_NEUCR	Q55154 synecocyst
75	31	57.4	788	1	P23622 neurospora

## ALIGNMENTS

## RESULT 1

CARB_FUSNN	STANDARD;	PRT; 1058 AA.
ID	Q8RG86;	
AC	15-JUN-2002 (Rel. 41, Created)	
DT	15-JUN-2002 (Rel. 41, Last sequence update)	
DT	15-JUN-2002 (Rel. 41, Last annotation update)	
DE	Carbamoyl-phosphatase large chain (EC 6.3.5.5) (Carbamoyl-phosphate synthase ammonia chain).	
DE	phosphatase synthase ammonia chain).	
GN	CARB OR FN0422.	
OS	Fusobacterium nucleatum (subsp. nucleatum).	
OC	Bacteria; Fusobacteria; Fusobacterium.	
OX	NCBI_TaxID=76856;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=ATCC 25586;	
RX	MEDLINE=21886394; PubMed=11889109;	
RA	Kapatal V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A.,	
RA	Bhattacharyya A., Bartman A., Gardner W., Grechkin G., Zhu L.,	
RA	Vasileva O., Chu L., Kogan Y., Chaga O., Goltzman E., Bernal A.,	
RA	Larsen N., D'Souza M., Walunas T., Pusck G., Haselkorn R.,	
RA	Fonstein M., Kyrplides N., Overbeek R.,	
RT	"Genome sequence and analysis of the oral bacterium Fusobacterium	
RT	nucleatum strain ATCC 25586."	
RL	J. Bacteriol. 184:2005-2018(2002).	
CC	- - CATALYTIC ACTIVITY: 2 ATP + L-glutamine + CO(2) + H(2)O = 2 ADP +	

phosphate + L-glutamate + carbamoyl phosphate.  
 -1- COFACTOR: Binds three manganese ions (By similarity).  
 -1- PATHWAY: Arginine biosynthesis.  
 -1- SUBUNIT: Composed of two chains; the small (or glutamine) chain promotes the hydrolysis of glutamine to ammonia, which is used by the large (or ammonia) chain to synthesize carbamoyl phosphate (By similarity).  
 -1- SIMILARITY: BELONGS TO THE CARB FAMILY.  
 -----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 EMBL: A6010554; AAL94625.1; ALT\_INIT.  
 DR InterPro: IPR005483; CPase\_L.  
 DR InterPro: IPR005479; CPase\_L2.  
 DR InterPro: IPR005480; CPase\_LD3.  
 DR InterPro: IPR005481; CPase\_LN.  
 DR InterPro: IPR004362; MGS-like.  
 DR Pfam: PF02789; CPase\_L\_chain; 2.  
 DR Pfam: PF02786; CPase\_L\_D2; 2.  
 DR Pfam: PF02787; CPase\_L\_D3; 1.  
 DR Pfam: PF02142; MGS; 1.  
 DR PRINTS: PR00098; CPASE.  
 DR PROSITE: PS00866; CPASE\_1; 2.  
 DR PROSITE: PS00867; CPASE\_2; 2.  
 KW Arginine biosynthesis; Pyrimidine biosynthesis; Ligase; Repeat;  
 KW ATP-binding; Manganese; Complete proteome.  
 FT DOMAIN 1 401 CARBOXYPHOSPHATE SYNTHETIC DOMAIN.  
 FT DOMAIN 402 546 OLIGOMERIZATION DOMAIN.  
 FT DOMAIN 547 929 CARBAMOYL PHOSPHATE SYNTHETIC DOMAIN.  
 FT REPEAT 930 1058 ALLOSTERIC DOMAIN.  
 FT REPEAT 1 546  
 FT REPEAT 547 1058  
 FT NP\_BIND 153 210 ATP (POTENTIAL).  
 FT NP\_BIND 302 352 ATP (POTENTIAL).  
 FT METAL 284 284 MANGANESE 1 (BY SIMILARITY).  
 FT METAL 298 298 MANGANESE 1 AND 2 (BY SIMILARITY).  
 FT METAL 300 300 MANGANESE 2 (BY SIMILARITY).  
 FT METAL 820 820 MANGANESE 3 (BY SIMILARITY).  
 FT METAL 832 832 MANGANESE 3 (BY SIMILARITY).  
 SQ SEQUENCE 1058 AA; 117451 MW; ED7037AF77C1E39F CRC64;  
 -----  
 Query Match 70.4%; Score 38; DB 1; Length 1058;  
 Best Local Similarity 60.0%; Pred. No. 7.8;  
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
 QY 2 EVVFXGMDYS 11  
 Db 190 EIVPGLNYS 199  
 :||| |::||  
 -----  
 RESULT 2  
 Y1LK\_TYDVA  
 ID Y1LK\_TYDVA STANDARD; PRT; 102 AA.  
 AC P31619;  
 DT 01-JUL-1993 (Rel. 26, Created)  
 DT 01-JUL-1993 (Rel. 26, Last sequence update)  
 DT 01-OCT-1993 (Rel. 27, Last annotation update)  
 DE Hypothetical 11.2 kDa protein.  
 GN Y1.  
 OS Tobacco yellow dwarf virus (strain Australia) (TYDV).  
 OC Viruses; ssDNA viruses; Geminiviridae; Mastrevirus.  
 OX NCBI\_TaxID=31599;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92188538; PubMed=1546450;  
 RA Morris B.A.M., Richardson K.A., Haley A., Zhan X., Thomas J.E.;

"The nucleotide sequence of the infectious cloned DNA component of tobacco yellow dwarf virus reveals features of geminiviruses infecting monocotyledonous plants.";  
 Virology 187:633-642(1992).  
 -----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 EMBL: M81103; AAA47947.1; -.  
 DR PIR: A42452; A42452.  
 DR InterPro: IPR002621; Gemini\_mov.  
 DR Pfam: PF01708; Gemini\_mov; 1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 102 AA; 11178 MW; A40ECF1E0AF55B67 CRC64;  
 -----  
 Query Match 66.7%; Score 36; DB 1; Length 102;  
 Best Local Similarity 60.0%; Pred. No. 1.7;  
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
 QY 2 EVVFXGMDYS 11  
 Db 7 QVPSGINS 16  
 :||| |::||  
 -----  
 RESULT 3  
 SPS\_ORYSA  
 ID SPS\_ORYSA STANDARD; PRT; 1049 AA.  
 AC Q43802;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE Sucrose-phosphate synthase (EC 2.4.1.14)  
 DE (UDP-glucose-fructose-phosphate glucosyltransferase).  
 OS Oryza sativa (Rice).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzoae; Oryza.  
 OX NCBI\_TaxID=4530;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Indica-IR36; TISSUE=Leaf;  
 RX MEDLINE=96235138; PubMed=8666248;  
 RA Valdez-Alarcon J.J., Ferrando M., Jimenez-Moralla B.,  
 RA Herrera-Estrella L.;  
 RT "Characterization of a rice sucrose-phosphate synthase-encoding gene.";  
 RT Gene 170:217-222(1996).  
 RL  
 CC -1- FUNCTION: INVOLVED IN THE REGULATION OF CARBON PARTITIONING IN THE LEAVES OF PLANTS. MAY REGULATE THE SYNTHESIS OF SUCROSE AND THEREFORE PLAY A MAJOR ROLE AS A LIMITING FACTOR IN THE EXPORT OF PHOASIMILATES OUT OF THE LEAF.  
 CC -1- CATALYTIC ACTIVITY: UDP-glucose + D-fructose 6-phosphate = UDP + sucrose 6-phosphate.  
 CC -1- ENZYME REGULATION: ACTIVITY REGULATED BY PHOSPHORYLATION AND MODERATED BY CONCENTRATION OF METABOLITES AND LIGHT.  
 CC -1- PATHWAY: Sucrose synthesis.  
 CC -1- SUBUNIT: HOMODIMER OR HOMOTETRAMER (BY SIMILARITY).  
 CC -1- PTM: PHOSPHORYLATED. HOWEVER, PHOSPHORYLATION IS NOT ESSENTIAL FOR ENZYME FUNCTION (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 1.  
 -----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).



```
CC EMBL; U33175; AAC49379.1; -
DR InterPro; IPR001296; Glycos_transf_1.
DR Pfam; PF00534; Glycosyltransferase; 1. Phosphorylation.
KW Transferase; Glycosyltransferase; 1. Phosphorylation.
FT DOMAIN 22 29 POLY-GLY.
FT DOMAIN 695 698 POLY-GLU.
FT DOMAIN 775 779 POLY-ARG.
SQ SEQUENCE 1049 AA; 116455 MW; ED862E2819AA4B04 CRC64;

Query Match 66.7%; Score 36; DB 1; Length 1049;
Best Local Similarity 66.7%; Pred. No. 20;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMDYS 11
DB 436 VIPPGMDFS 444

RESULT 4.
SPS_MAIZE
ID SPS_MAIZE STANDARD; PRT; 1068 AA.
AC P31927;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE Sucrose-phosphate synthase 2 (EC 2.4.1.14) (UDP-glucose-fructose-
phosphate glucosyltransferase).
GN SPS.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC Clade;
OC Panicoidae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 71-74; 206-212; 471-481 AND
RP 872-892.
RC STRAIN=CV. PIONEER 3184; TISSUE=Leaf;
RX MEDLINE=92338837; PubMed=1840396;
RA Worrell A.C., Bruneau J.-M., Summerfelt K., Boersig M., Voelker T.A.;
RT "Expression of a maize sucrose phosphate synthase in tomato alters
RT leaf carbohydrate partitioning.";
RL Plant Cell 3:1121-1130(1991).
CC -1- FUNCTION: INVOLVED IN THE REGULATION OF CARBON PARTITIONING IN
CC THE LEAVES OF PLANTS. MAY REGULATE THE SYNTHESIS OF SUCROSE AND
CC THEREFORE PLAY A MAJOR ROLE AS A LIMITING FACTOR IN THE EXPORT OF
CC PHOTOASSIMILATES OUT OF THE LEAF.
CC -1- CATALYTIC ACTIVITY: UDP-glucose + D-fructose 6-phosphate = UDP +
CC sucrose 6-phosphate.
CC -1- ENZYME REGULATION: ACTIVITY REGULATED BY PHOSPHORYLATION AND
CC MODERATED BY CONCENTRATION OF METABOLITES AND LIGHT.
CC -1- PATHWAY: Sucrose synthesis.
CC -1- SUBUNIT: HOMODIMER OR HOMOTETRAMER.
CC -1- DEVELOPMENTAL STAGE: GERMINATING SEEDS OR MATURE LEAVES.
CC -1- PTM: PHOSPHORYLATED. HOWEVER, PHOSPHORYLATION IS NOT ESSENTIAL FOR
CC ENZYME FUNCTION.
CC -1- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 1.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics. and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M97550; AAA33513.1; -
DR PIR; JQ1329; JQ1329.
DR MaizeDB; 25294; -
DR InterPro; IPR001296; Glycos_transf_1.
DR Pfam; PF00534; Glycosyltransferase; 1.
KW Transferase; Glycosyltransferase; Phosphorylation.
FT DOMAIN 25 31 POLY-GLY.
FT
```

```
SQ SEQUENCE 1068 AA; 118575 MW; 074679B5E9A1D282 CRC64;

Query Match 66.7%; Score 36; DB 1; Length 1068;
Best Local Similarity 66.7%; Pred. No. 21;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMDYS 11
DB 435 VIPPGMDFS 443

RESULT 5.
SPS2_CRAPL
ID SPS2_CRAPL STANDARD; PRT; 1081 AA.
AC O04933;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Sucrose-phosphate synthase 2 (EC 2.4.1.14) (UDP-glucose-fructose-
phosphate glucosyltransferase 2).
GN SPS2.
OS Craterostigma plantaginum.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Toreneae;
OC Asteridae; euasterids I; Lamiales; Lamiales incertae sedis; Toreneae;
OC Craterostigma.
OX NCBI_TaxID=4153;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97451773; PubMed=9306694;
RA Ingram J., Chandler J.W., Gallagher L., Salami F., Bartels D.;
RT "Analysis of cDNA clones encoding sucrose-phosphate synthase in
RT relation to sugar interconversions associated with dehydration in the
RT resurrection plant Craterostigma plantaginum Hochst.";
RL Plant Physiol. 115:113-121(1997).
CC -1- FUNCTION: INVOLVED IN THE REGULATION OF CARBON PARTITIONING IN
CC THE LEAVES OF PLANTS. MAY REGULATE THE SYNTHESIS OF SUCROSE AND
CC THEREFORE PLAY A MAJOR ROLE AS A LIMITING FACTOR IN THE EXPORT OF
CC PHOTOASSIMILATES OUT OF THE LEAF.
CC -1- CATALYTIC ACTIVITY: UDP-glucose + D-fructose 6-phosphate = UDP +
CC sucrose 6-phosphate.
CC -1- ENZYME REGULATION: ACTIVITY REGULATED BY PHOSPHORYLATION AND
CC MODERATED BY CONCENTRATION OF METABOLITES AND LIGHT.
CC -1- PATHWAY: Sucrose synthesis.
CC -1- SUBUNIT: HOMODIMER OR HOMOTETRAMER (BY SIMILARITY).
CC -1- PTM: PHOSPHORYLATED. HOWEVER, PHOSPHORYLATION IS NOT ESSENTIAL FOR
CC ENZYME FUNCTION (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 1.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Y11795; CAA72491.1; -
DR InterPro; IPR001296; Glycos_transf_1.
DR Pfam; PF00534; Glycosyltransferase; 1.
KW Transferase; Glycosyltransferase; Phosphorylation; Multigene family.
FT DOMAIN 245 248 POLY-SER.
FT DOMAIN 256 264 POLY-GLU.
FT DOMAIN 787 790 POLY-ARG.
SQ SEQUENCE 1081 AA; 120933 MW; DD142DC2F1A72900 CRC64;

Query Match 66.7%; Score 36; DB 1; Length 1081;
Best Local Similarity 66.7%; Pred. No. 21;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMDYS 11
DB 445 VIPPGMDFS 453
```

RESULT 6  
FAF\_DROME  
ID FAF\_DROME STANDARD; PRT; 2747 AA.  
AC P55824;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 16-OCT-2001 (Rel. 35, Last sequence update)  
DE Probable ubiquitin carboxyl-terminal hydrolase FAF (EC 3.1.2.15)  
DE (Ubiquitin thiolesterase FAF) (Ubiquitin-specific processing protease FAF)  
GN FAF (Deubiquitinating enzyme FAF) (Fat facets protein).  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;  
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;  
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93202020; PubMed=1295747;  
RA Fischer-Vize J.A., Rubin G.M., Lehmann R.;  
RT "The fat facets gene is required for Drosophila eye and embryo development.";  
RL Development 116:985-1000(1992).  
CC -1- FUNCTION: REQUIRED FOR EYE AND EMBRYO DEVELOPMENT, AND PLAYS A  
CC ROLE IN COMPOUND EYE ASSEMBLY AND OGENESIS RESPECTIVELY. IN THE  
CC LARVAL EYE DISKS, CELLS OUTSIDE THE ASSEMBLING FACETS REQUIRE THIS  
CC PROTEIN FOR SHORT-RANGE CELL INTERACTIONS THAT PREVENT THE MYSTERY  
CC CELLS FROM BECOMING PHOTORECEPTORS. IT IS ALSO REQUIRED FOR  
CC NUCLEAR MIGRATION AND CELLULARIZATION IN EARLY EMBRYOGENESIS AND  
CC COULD PLAY A ROLE IN POLE CELL DETERMINATION, DEVELOPMENT OR  
CC FUNCTION.  
CC -1- CATALYTIC ACTIVITY: Ubiquitin C-terminal thiolester + H(2)O =  
CC ubiquitin + a thiol.  
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A  
CC SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.  
CC -1- TISSUE SPECIFICITY: EYE DISKS AND OVARIES.  
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C19.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; L04959; AAF01345.1; -;  
CC EMBL; L04958; AAF01346.1; -;  
CC MEROPS; C19.007; -;  
CC FlyBase; FBgn0005632; faf.  
CC InterPro; IPR001394; UCH-2.  
CC Pfam; PF00442; UCH-1; 1.  
CC Pfam; PF00443; UCH-2; 1.  
CC PROSITE; PS00972; UCH\_2\_1; 1.  
CC PROSITE; PS00973; UCH\_2\_2; 1.  
CC PROSITE; PS02335; UCH\_2\_3; 1.  
CC Ubi conjugation pathway; Hydrolase; Thiol protease;  
KW Developmental protein; Vision; Alternative splicing.  
FT ACT\_SITE 1677 1677 BY SIMILARITY.  
FT ACT\_SITE 1978 1978 BY SIMILARITY.  
FT ACT\_SITE 1986 1986 BY SIMILARITY.  
FT VARSPPLIC 2705 2747 KCRVLIKKLVESKDEEDATSAATTAATTEVTTSPATAS  
FT VARIANT 2725 2725 S -> T.  
FT SEQUENCE 2747 AA; 307954 MW; ID97659F7A7B2ADE CRC64;  
Query Match 64.8%; Score 35; DB 1; Length 2747;  
Best Local Similarity 54.5%; Pred. No. 91;  
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
QY 1 BEVVPXGMDYS 11

Db 1394 EVIVPDGQDFS 1404  
RESULT 7  
KHL1\_HUMAN  
ID KHL1\_HUMAN STANDARD; PRT; 748 AA.  
AC Q9NR64; Q9NR65; Q9P238; Q9H4X4;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DE (Ubiquitin thiolesterase FAF) (Ubiquitin-specific processing protease FAF)  
GN FAF (Deubiquitinating enzyme FAF) (Fat facets protein).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20347694; PubMed=1088605;  
RA Koob M.D., Nemes J.P., Benzow K.A.;  
RT "The SCAB transcript is an antisense RNA to a brain-specific  
RT transcript encoding a novel actin-binding protein (KLHL1).";  
RL Hum. Mol. Genet. 9:1543-1551(2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RX MEDLINE=20277482; PubMed=10819331;  
RA Nagaya T., Kikuno R., Ishikawa K.-I., Hirose M., Ohara O.;  
RT "Prediction of the coding sequences of unidentified human genes. XVII.  
RT the complete sequences of 100 new cDNA clones from brain which code  
RT for large proteins in vitro.";  
RL DNA Res. 7:143-150(2000).  
RN [3]  
RP SEQUENCE OF 179-409 FROM N.A.  
RA Kay M.;  
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: MAY PLAY A ROLE IN ORGANIZING THE ACTIN CYTOSKELETON OF  
CC THE BRAIN CELLS.  
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN BRAIN.  
CC -1- SIMILARITY: CONTAINS 1 BTB/POZ DOMAIN.  
CC -1- SIMILARITY: CONTAINS 6 KELCH REPEATS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; AF252283; AAF81719.1; -;  
CC EMBL; AF252279; AAF81716.1; -;  
CC EMBL; AB040923; BAA96014.1; ALT\_INIT.  
CC EMBL; AL353738; CAC16128.1; -;  
CC Genew; HGNC:6352; KLHL1.  
CC MIM; 605332; -;  
CC InterPro; IPR000210; BTB\_POZ.  
CC InterPro; IPR001798; Kelch.  
CC Pfam; PF00651; BTB; 1.  
CC Pfam; PF01344; Kelch; 6.  
CC PRINTS; PR00501; KELCHREPEAT.  
CC SMART; SM00225; BTB; 1.  
CC PROSITE; PS50097; BTB; 1.  
CC Cytoskeleton; Actin-binding; Repeat.  
KW DOMAIN 43 88  
FT DOMAIN 212 279 BTB.  
FT REPEAT 460 506 KELCH 1.  
FT REPEAT 507 553 KELCH 2.  
FT REPEAT 555 600 KELCH 3.  
FT REPEAT 601 647 KELCH 4.  
FT REPEAT 649 700 KELCH 5.

FT	REPEAT	701	747	KELCH 6.	
SQ	SEQUENCE	748 AA;	82680 MW;	CL1C4D8282F9FF9 CRC64;	
	Query Match	63.9%;	Score 34.5;	DB 1;	Length 748;
	Best Local Similarity	80.0%;	Pred. No. 29;		
	Matches	8;	Conservative	1;	Mismatches 0;
				Indels	1;
				Gaps	1;
QY	1	EEVVPXGMDY 10			
Db	127	EEVVP-GMDF 135			
RESULT 8					
ID	PLAS_ORYSA	STANDARD;	PRT;	154 AA.	
AC	P20423; Q9SBB8;				
DT	01-FEB-1991 (Rel. 17, Created)				
DT	15-JUN-2002 (Rel. 41, Last sequence update)				
DT	15-JUN-2002 (Rel. 41, Last annotation update)				
DE	Plastocyanin, chloroplast precursor.				
GN	PETE.				
OS	Oryza sativa (Rice).				
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;				
OC	Eurhartoideae; Oryzeae; Oryza.				
OX	NCBI_TaxID=4530;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=cv. Ilpoom; TISSUE=Leaf;				
RA	Lee J.-S.;				
RT	"Molecular cloning and characterization of plastocyanin precursor in rice.";				
RL	Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.				
RN	[2]				
RP	SEQUENCE OF 58-154.				
RC	STRAIN=cv. Japonica;				
RX	MEDLINE=89386623; PubMed=2780537;				
RA	Yano H., Kamo M., Tsugita A., Aso K., Nozu Y.;				
RT	"The amino acid sequence of plastocyanin from rice (Oryza sativa, subspecies japonica).";				
RL	Protein Seq. Data Anal. 2:385-389(1989).				
CC	-1- FUNCTION: Participates in electron transfer between P700 and the cytochrome b6-f complex in photosystem I.				
CC	-1- SUBCELLULAR LOCATION: LOOSELY BOUND TO THE INNER THYLAKOID MEMBRANE SURFACE IN CHLOROPLASTS.				
CC	-1- SIMILARITY: CONTAINS 1 PLASTOCYANIN-LIKE DOMAIN.				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).				
CC	EMBL; AF093636; AAC78108.1;				
DR	PIR; S06105; S06105.				
DR	PIR; J03552; J03552.				
DR	HSSP; P00289; 2PCF.				
DR	InterPro; IPR000923; BlueCu_1.				
DR	InterPro; IPR001235; Copper_blue.				
DR	Pfam; PF00127; copper-bind; 1.				
DR	PRINTS; PR00156; COPPERBLUE.				
DR	ProDom; PD001235; Copper_blue; 1.				
DR	PROSITE; PS00196; COPPER_BLUE; 1.				
KW	Chloroplast; Electron transport; Copper; Thylakoid; Membrane;				
KW	Transit peptide.				
FT	TRANSIT	1	57	CHLOROPLAST.	
FT	CHAIN	58	154	PLASTOCYANIN.	
FT	DOMAIN	58	154	PLASTOCYANIN-LIKE.	
FT	METAL	94	94	COPPER (BY SIMILARITY).	
FT	METAL	139	139	COPPER (BY SIMILARITY).	
FT	METAL	142	142	COPPER (BY SIMILARITY).	

FT	METAL	147	147	COPPER (BY SIMILARITY).	
SQ	SEQUENCE	154 AA;	15577 MW;	E45725D25B5F400D CRC64;	
	Query Match	63.0%;	Score 34;	DB 1;	Length 154;
	Best Local Similarity	54.5%;	Pred. No. 6.8;		
	Matches	6;	Conservative	2;	Mismatches 3;
				Indels	0;
				Gaps	0;
QY	1	EEVVPXGMDYS 11			
		:       :			
Db	100	EDAVPSGVDVS 110			
RESULT 9					
ID	PLAS_HORVU	STANDARD;	PRT;	155 AA.	
AC	P08248;				
DT	01-AUG-1988 (Rel. 08, Created)				
DT	01-AUG-1988 (Rel. 08, Last sequence update)				
DT	15-JUN-2002 (Rel. 41, Last annotation update)				
DE	Plastocyanin, chloroplast precursor.				
GN	PETE.				
OS	Hordeum vulgare (Barley).				
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;				
OC	Triticeae; Hordeum.				
OX	NCBI_TaxID=4513;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=cv. Bomi;				
RA	Nielsen O.S., Gausung K.;				
RT	"The precursor of barley plastocyanin: sequence of cDNA clones and gene expression in different tissues.";				
RL	FEBS Lett. 225:159-162(1987).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=cv. NK 1558;				
RX	MEDLINE=94039081; PubMed=8223592;				
RA	Nielsen P., Gausung K.;				
RT	"In vitro binding of nuclear proteins to the barley plastocyanin gene promoter region.";				
RL	Eur. J. Biochem. 217:97-104(1993).				
CC	-1- FUNCTION: Participates in electron transfer between P700 and the cytochrome b6-f complex in photosystem I.				
CC	-1- SUBCELLULAR LOCATION: LOOSELY BOUND TO THE INNER THYLAKOID MEMBRANE SURFACE IN CHLOROPLASTS.				
CC	-1- SIMILARITY: CONTAINS 1 PLASTOCYANIN-LIKE DOMAIN.				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).				
CC	EMBL; Y00704; CAA68696.1;				
DR	EMBL; Z28347; CAA82201.1;				
DR	PIR; S00206; S00206.				
DR	HSSP; P00289; 2PCF.				
DR	InterPro; IPR000923; BlueCu_1.				
DR	InterPro; IPR001235; Copper_blue.				
DR	Pfam; PF00127; copper-bind; 1.				
DR	PRINTS; PR00156; COPPERBLUE.				
DR	ProDom; PD001235; Copper_blue; 1.				
DR	PROSITE; PS00196; COPPER_BLUE; 1.				
KW	Chloroplast; Electron transport; Copper; Thylakoid; Membrane;				
KW	Transit peptide.				
FT	TRANSIT	1	58	CHLOROPLAST.	
FT	CHAIN	59	155	PLASTOCYANIN.	
FT	DOMAIN	59	155	PLASTOCYANIN-LIKE.	
FT	METAL	95	95	COPPER (BY SIMILARITY).	
FT	METAL	140	140	COPPER (BY SIMILARITY).	
FT	METAL	143	143	COPPER (BY SIMILARITY).	

FT METAL 148 148 COPPER (BY SIMILARITY).  
 FT VARIANT 120 120 T -> N (IN CV. NK 1558).  
 SQ SEQUENCE 155 AA; 15709 MW; DAA7EABE5F6F4F91 CRC64;

Query Match 63.0%; Score 34; DB 1; Length 155;  
 Best Local Similarity 54.5%; Pred. No. 6.9;  
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMDYS 11  
 I: ||| |  
 Db 101 EDVPSGVDS 111

## RESULT 10

PLAT\_POPNI STANDARD; PRT; 168 AA.  
 AC P11970;  
 DT 01-OCT-1989 (Rel. 12, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Plastocyanin B, chloroplast precursor.  
 GN PTE.  
 OS Populus nigra (Lombardy poplar).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids I; Malpighiales; Salicaceae; Populus.  
 OX NCBI\_TaxID=3691;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Italica; TISSUE=Leaf;  
 RA Reichert J., Jenzelewski V., Haeihel W.;  
 RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE OF 70-168.  
 RC STRAIN=cv. Italica;  
 RA Dmitrov M.I., Egorov C.A., Donchev A.A., Atanasov B.P.;  
 RT "Complete amino acid sequence of poplar plastocyanin b.";  
 RL FEBS Lett. 226:17-22(1987).  
 CC -1- FUNCTION: Participates in electron transfer between P700 and the  
 CC cytochrome b6-f complex in photosystem I.  
 CC -1- SUBCELLULAR LOCATION: LOOSELY BOUND TO THE INNER THYLAKOID  
 CC MEMBRANE SURFACE IN CHLOROPLASTS.  
 CC -1- SIMILARITY: THERE ARE 12 DIFFERENCES BETWEEN THE SEQUENCE OF  
 CC POPULAR PLASTOCYANINS A AND B.  
 CC -1- SIMILARITY: CONTAINS 1 PLASTOCYANIN-LIKE DOMAIN.

-----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----

EMBL; Z50186; CAA90565.1; -;  
 DR PIR; S00210; S00210.  
 DR HSP; P00299; IPLC.  
 DR InterPro; IPR000923; BlueCu.1.  
 DR InterPro; IPR001235; Copper.blue.  
 DR Pfam; PF00127; copper-bind; 1.  
 DR PRINTS; PR00156; COPPERBLUE.  
 DR ProDom; PD001235; Copper\_blue; 1.  
 DR PROSITE; PS00196; COPPER\_BLUE; 1.  
 KW Chloroplast; Electron transport; Copper; Thylakoid; Membrane;  
 KW Transit peptide; Multigene family.  
 FT TRANSIT 1 69 CHLOROPLAST.  
 FT CHAIN 70 168 PLASTOCYANIN B.  
 FT DOMAIN 70 168 PLASTOCYANIN-LIKE.  
 FT METAL 106 106 COPPER.  
 FT METAL 153 153 COPPER.  
 FT METAL 156 156 COPPER.  
 FT METAL 161 161 COPPER.  
 SQ SEQUENCE 168 AA; 16981 MW; F20DA6EA2038AE6A CRC64;

Query Match 63.0%; Score 34; DB 1; Length 168;  
 Best Local Similarity 54.5%; Pred. No. 7.5;  
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMDYS 11  
 I: ||| |  
 Db 112 EDVPSGVDS 122

## RESULT 11

SYFB\_BORBU STANDARD; PRT; 566 AA.  
 AC P94283;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Phenylalanyl-tRNA synthetase beta chain (EC 6.1.1.20) (Phenylalanine--  
 DE tRNA ligase beta chain) (PHERS).  
 GN PHE1 OR BB0514.  
 OS Borrelia burgdorferi (Lyme disease spirochete).  
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
 OX NCBI\_TaxID=139;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 35210 / B31;  
 RA Barbour A.G., Hinnebusch J.;  
 RT "Phenylalanyl-tRNA synthetase genes (alpha and beta subunits) and  
 RT thioredoxin reductase gene of Borrelia burgdorferi";  
 RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 35210 / B31;  
 RX MEDLINE=98065943; PubMed=9403685;  
 RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,  
 RA Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,  
 RA Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,  
 RA Peterson J., Kleravage A.R., Quackenbush J., Salzberg S., Hanson M.,  
 RA van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,  
 RA Uterback T., Watthey L., McDonald L., Artiach P., Bowman C.,  
 RA Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,  
 RA Smith H.O., Venter J.C.;  
 RT "Genomic sequence of a Lyme disease spirochete, Borrelia  
 RT burgdorferi";  
 RL Nature 390:580-586(1997).  
 CC -1- CATALYTIC ACTIVITY: ATP + L-phenylalanine + tRNA(Phe) = AMP +  
 CC diphosphate + L-phenylalanyl-tRNA(Phe).  
 CC -1- SUBUNIT: Tetramer of two alpha and two beta chains (BY  
 CC SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -1- SIMILARITY: BELONGS TO THE PHENYLALANYL-TRNA SYNTHETASE BETA  
 CC CHAIN FAMILY. SUBFAMILY 2.  
 -----

This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----

EMBL; U82978; AAB41019.1; -;  
 DR EMBL; AE001153; AAC66870.1; -;  
 DR TIGR; BB0514; -;  
 DR InterPro; IPR005147; B5.  
 DR InterPro; IPR004531; PheT\_arch.  
 DR Pfam; PF03484; B5; 1.  
 DR TIGRFAMs; TIGR00471; pheT\_arch; 1.  
 KW Aminoacyl-tRNA synthetase; Ligase; ATP-binding;  
 KW Complete proteome.  
 SQ SEQUENCE 566 AA; 65173 MW; 9D48C8B5D6D3B74B CRC64;

Query Match 63.0%; Score 34; DB 1; Length 566;

```

Best Local Similarity 85.7%; Pred. No. 27;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 VPXGMDY 10
DB 169 VPFGMDY 175

RESULT 12
ID Y939_METJA STANDARD; PRT; 276 AA.
AC Q58349;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein MJ0939.
GN MJ0939.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
OX NCBI_TaxID=2190;
RN SEQUENCE FROM N.A.
RP STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghagen N.S.M., Weidman J.F., Fuhrman J.L., Nguyen D.,
RA Uitterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii.";
RL Science 273:1058-1073(1996).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: U67537; AAB98946.1; -.
DR TIGR: MJ0939; -.
DR TIGR: MJ0939; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 276 AA; 33454 MW; 97BD69D392BC8FDF CRC64;

Query Match 61.1%; Score 33; DB 1; Length 276;
Best Local Similarity 45.5%; Pred. No. 21;
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMDYS 11
DB 141 EEIENGMEHS 151

RESULT 13
ID PANC_PSEAE STANDARD; PRT; 283 AA.
AC Q9HV69;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Pantoate-beta-alanine ligase (EC 6.3.2.1) (Pantothenate synthetase)
DE (Pantoate activating enzyme).
GN PANC OR PA4730.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=287;

[1]
RN SEQUENCE FROM N.A.
RP STRAIN=ATCC 15692 / PA01;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrenner P.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Coulter L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen.";
RL Nature 406:959-964(2000).
CC -!- CATALYTIC ACTIVITY: ATP + (R)-pantoate + beta-alanine = AMP +
CC diphosphate + (R)-pantothenate.
CC -!- PATHWAY: Pantothenate biosynthesis; last step.
CC -!- SIMILARITY: BELONGS TO THE PANTOTHENATE SYNTHETASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AE004886; AAG08116.1; -.
DR InterPro: IPR003721; Pantoate_ligase.
DR Pfam: PF02569; Pantoate_ligase; 1.
DR TIGR: TIGR00018; pncC; 1.
KW Pantothenate biosynthesis; Ligase; Complete proteome.
SQ SEQUENCE 283 AA; 30836 MW; C494949AB40E14E7 CRC64;

Query Match 61.1%; Score 33; DB 1; Length 283;
Best Local Similarity 66.7%; Pred. No. 21;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMD 9
DB 96 EEMYPDGM 104

RESULT 14
ID HMPA_VIBCH STANDARD; PRT; 394 AA.
AC Q9KMY3;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Flavohemoprotein (Hemoglobin-like protein) (Flavohemoglobin).
GN HMP OR VCA0183.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;

[1]
RN SEQUENCE FROM N.A.
RP STRAIN=El Tor N16961 / Serotype O1;
RX MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
RA McDonald L., Uitterback T., Fleischmann R.D., Nierman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae.";
RL Nature 406:477-483(2000).
CC -!- DOMAIN: CONSISTS OF TWO DISTINCT DOMAINS; ONE IS A HEME-CONTAINING
CC OXYGEN BINDING DOMAIN IN THE N-TERMINAL REGION AND THE OTHER IS AN
CC FAD-CONTAINING REDUCTASE DOMAIN FOUND IN THE C-TERMINAL REGION.
CC -!- SIMILARITY: BELONGS TO THE GLOBIN FAMILY. TWO-DOMAIN
CC FLAVOHEMOPROTEINS SUBFAMILY.

```

```

CC CC -!- SIMILARITY: TO A NUMBER OF FAD/NAD(P) FLAVOPROTEIN
CC CC OXIDOREDUCTASES.
CC CC
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
CC CC EMBL; AE004358; AAF96096.1; -
CC CC HSSP; P39662; 1CQX.
CC CC TIGR; VCA0183; -
CC CC InterPro; IPR001834; Cyt_B5_reductase.
CC CC InterPro; IPR001709; FPN_cyt_redctse.
CC CC InterPro; IPR000971; Globin.
CC CC InterPro; IPR001433; Oxred_FAD/NAD(P).
CC CC InterPro; IPR001221; Phe_hydroxylase.
CC CC Pfam; PF00042; globin; 1.
CC CC Pfam; PF00175; NAD_binding; 1.
CC CC Pfam; PF00970; FAD_binding_6; 1.
CC CC PRINTS; PR00371; FPNCR.
CC CC PRINTS; PR00410; PHEHYDRLASE.
CC CC PROSITE; PS01033; GLOBIN; 1.
CC CC Oxidoreductase; NADP; Heme; Flavoprotein; FAD; Iron transport;
CC CC Oxygen transport; Transport; Complete proteome.
CC CC FT DOMAIN 1 136
CC CC FT METAL 53 53
CC CC FT METAL 85 85
CC CC FT METAL 268 273
CC CC NP_BIND 268 273
CC CC NADP (RIBOSE PART) (BY SIMILARITY).
CC CC SEQUENCE 394 AA; 44191 MW; DDA3490FAE28823A CRC64;
CC CC
CC CC Query Match 61.1%; Score 33; DB 1; Length 394;
CC CC Best Local Similarity 66.7%; Pred. No. 30;
CC CC Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
CC CC
CC CC QY 2 EVVPGXGMDY 10
CC CC |||||
CC CC Db 194 EVTPEGSDY 202
CC CC
CC CC RESULT 15
CC CC ACADM_RAT
CC CC ID ACADM_RAT STANDARD; PRT; 421 AA.
CC CC AC P08503;
CC CC DT 01-AUG-1988 (Rel. 08, Created)
CC CC DT 01-AUG-1988 (Rel. 08, Last sequence update)
CC CC DT 15-JUN-2002 (Rel. 41, Last annotation update)
CC CC DE Acyl-CoA dehydrogenase, medium-chain specific, mitochondrial precursor
CC CC (EC 1.3.99.3) (NCAD).
CC CC GN ACADM.
CC CC OS Rattus norvegicus (Rat).
CC CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC CC OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
CC CC OX NCBI_TaxID=10116;
CC CC RN [1]
CC CC RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
CC CC RC TISSUE=Liver;
CC CC RX MEDLINE=87280028; PubMed=3611054;
CC CC RA Matsubara Y., Kraus J.P., Ozasa H., Glassberg R., Finocchiaro G.,
CC CC Ikeda Y., Mole J., Rosenberg L.E., Tanaka K.;
CC CC RT "Molecular cloning and nucleotide sequence of cDNA encoding the
CC CC entire precursor of rat liver medium chain acyl coenzyme A
CC CC dehydrogenase.";
CC CC RL J. Biol. Chem. 262:10104-10108(1987).
CC CC -!- FUNCTION: 262:10104-10108(1987).
CC CC -!- CATALYTIC ACTIVITY: Acyl-CoA + ETF = 2,3-dehydroacyl-CoA + reduced
CC CC ETF.
CC CC -!- COFACTOR: FAD.

```

```

CC CC -!- PATHWAY: Mitochondrial fatty acid beta-oxidation system; first
CC CC step.
CC CC -!- SUBUNIT: HOMOTETRAMER.
CC CC -!- MISCELLANEOUS: A NUMBER OF STRAIGHT-CHAIN ACYL-COA DEHYDROGENASES
CC CC OF DIFFERENT SUBSTRATE SPECIFICITIES ARE PRESENT IN MAMMALIAN
CC CC TISSUES.
CC CC -!- SIMILARITY: BELONGS TO THE ACYL-COA DEHYDROGENASE FAMILY.
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
CC CC EMBL; J02791; AAA40670.1; -
CC CC PIR; A28436; DERTCM.
CC CC HSSP; P11310; 1EGD.
CC CC InterPro; IPR001552; Acyl-CoA_dh.
CC CC Pfam; PF00441; Acyl-CoA_dh_1.
CC CC Pfam; PF02770; Acyl-CoA_dh_M; 1.
CC CC Pfam; PF02771; Acyl-CoA_dh_N; 1.
CC CC PROSITE; PS00072; ACYL_COA_DH_1; 1.
CC CC PROSITE; PS00073; ACYL_COA_DH_2; 1.
CC CC Oxidoreductase; Flavoprotein; FAD; fatty acid metabolism;
CC CC Mitochondrion; Transit peptide.
CC CC FT TRANSIT 1 25
CC CC FT CHAIN 26 421
CC CC ACYL-COA DEHYDROGENASE, MEDIUM-CHAIN
CC CC SPECIFIC.
CC CC ACT_SITE 193 193
CC CC FORMS A HYDROGEN-BOND WITH THE FLAVIN
CC CC N(5) OF THE FAD COFACTOR (BY SIMILARITY).
CC CC ACT_SITE 401 401
CC CC BASE (BY SIMILARITY).
CC CC SEQUENCE 421 AA; 46555 MW; 2CF076F8C919DE8 CRC64;
CC CC
CC CC Query Match 61.1%; Score 33; DB 1; Length 421;
CC CC Best Local Similarity 50.0%; Pred. No. 32;
CC CC Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
CC CC
CC CC QY 1 EEVVPXGMDY 10
CC CC |||||
CC CC Db 58 EEIIPVADPY 67
CC CC
CC CC RESULT 16
CC CC ECB2_HALEL
CC CC ID ECB2_HALEL STANDARD; PRT; 421 AA.
CC CC AC O52250;
CC CC DT 30-MAY-2000 (Rel. 39, Created)
CC CC DT 30-MAY-2000 (Rel. 39, Last sequence update)
CC CC DT 16-OCT-2001 (Rel. 40, Last annotation update)
CC CC DE Diaminobutyrate-pyruvate aminotransferase (EC 2.6.1.46) (L-
CC CC diaminobutyric acid transaminase) (Diaminobutyrate transaminase).
CC CC GN ECB.
CC CC OS Halomonas elongata.
CC CC OC Bacteria; Proteobacteria; gamma subdivision; Halomonadaceae;
CC CC OC Halomonas.
CC CC OX NCBI_TaxID=2746;
CC CC RN [1]
CC CC RP SEQUENCE FROM N.A.
CC CC RC STRAIN=DSM 2581T;
CC CC RX MEDLINE=98231640; PubMed=9570121;
CC CC RA Goller K., Ofer A., Galinski E.A.;
CC CC RT "Construction and characterization of an NaCl-sensitive mutant of
CC CC Halomonas elongata impaired in ectoine biosynthesis.";
CC CC RL FEMS Microbiol. Lett. 161:293-300(1998).
CC CC -!- CATALYTIC ACTIVITY: L-2,4-diaminobutanoate + pyruvate = L-
CC CC aspartate 4-semialdehyde + L-alanine.
CC CC -!- COFACTOR: PYRIDOXAL PHOSPHATE.
CC CC -!- PATHWAY: FIRST STEP IN THE BIOSYNTHESIS OF ECTOINE (1,4,5,6-
CC CC TETRAHYDRO-2-METHYL-4-PYRIMIDINE CARBOXYLIC ACID).
CC CC -!- SUBUNIT: HOMOHXAMER (BY SIMILARITY).

```

DR InterPro: IPR000954; AminoTran\_3.  
DR InterPro: IPR004637; Dat.  
DR Pfam: PF00202; aminoTran\_3; 1.  
DR TIGRFAMs: TIGR00709; dat; 1.  
DR PROSITE: PS00600; AA\_TRANSFER\_CLASS\_3; FALSE\_NEG.  
KW Transferase; AminoTransferase; PYRIDOXAL PHOSPHATE.  
FT BINDING 267 267 PYRIDOXAL PHOSPHATE (POTENTIAL).  
SQ SEQUENCE 423 AA; 46200 MW; 735C6BCF5A88288C CRC64;  
  
Query Match 61.1%; Score 33; DB 1; Length 423;  
Best Local Similarity 58.3%; Pred. No. 32;  
Matches 7; Conservative 2; Mismatches 1; Indels 2; Gaps 1;  
  
QY 1 EEVV--PXGMDY 10  
| | | | | | | | | |  
Db 91 EEVILKPRGLDY 102  
  
RESULT 18  
ECE2\_HUMAN STANDARD; PRT; 787 AA.  
ID ECE2\_HUMAN STANDARD; PRT; 787 AA.  
AC 060344; Q96NX4; Q96NX3;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 15-JUN-2002 (Rel. 41, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Endothelin-converting enzyme 2 (EC 3.4.24.71) (ECE-2).  
GN ECE2 OR KIAA0604.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORMS A; B AND C).  
RA Lorenzen M.-N., Khan R.Y., Wang Y., Tai S.C., Chan G.C., Cheung A.H.,  
RA Marsden P.A.;  
RT "Human endothelin converting enzyme-2 (ECE2): characterization of mRNA  
species and chromosomal localization.";  
RL Biochim. Biophys. Acta 0:0-0(2002).  
RN [2]  
RP SEQUENCE FROM N.A. (ISOFORM B).  
RT TISSUE=Brain;  
RX MEDLINE=98290545; PubMed=9628581;  
RA Nagase T., Ishikawa K.-I., Miyajima N., Tanaka A., Kotani H.,  
RA Nomura N., Ohara O.;  
RT "Prediction of the coding sequences of unidentified human genes. IX.  
The complete sequences of 100 new cDNA clones from brain which can  
code for large proteins in vitro.";  
RL DNA Res. 5:31-39(1998).  
CC -!- FUNCTION: CONVERTS BIG ENDOTHELIN-1 TO ENDOTHELIN-1 (BY  
SIMILARITY).  
CC -!- CATALYTIC ACTIVITY: Forms endothelin 1 by cleavage of the 21-  
Tyr-1-Val-22 bond in the precursor.  
CC -!- SUBCELLULAR LOCATION: Type II membrane protein.  
CC -!- ALTERNATIVE PRODUCTS: 3 isoforms; ECE-2A (shown here), ECE-2B and  
ECE-2C; are produced by alternative splicing.  
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M13.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC EMBL: AF428263; AAL30386.1; -  
CC EMBL: AF428264; AAL30387.1; -  
CC EMBL: AF192531; AAG28399.1; -  
CC EMBL: AB011176; BAA25530.1; -  
CC HSP: P08473; 1DMT.  
CC MEROPS: M13.003; -  
DR InterPro: IPR000718; Peptidase\_M13.  
DR InterPro: IPR000130; Zn\_MTPeptidse.  
DR

CC -!- SIMILARITY: BELONGS TO CLASS-III OF PYRIDOXAL-PHOSPHATE-DEPENDENT  
CC AMINOTRANSFERASES.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: AF031489; AAC15882.1; -  
CC HSP: P12995; 1QJ3.  
DR InterPro: IPR000954; AminoTran\_3.  
DR InterPro: IPR004637; Dat.  
DR Pfam: PF00202; aminoTran\_3; 1.  
DR TIGRFAMs: TIGR00709; dat; 1.  
DR PROSITE: PS00600; AA\_TRANSFER\_CLASS\_3; FALSE\_NEG.  
KW Transferase; AminoTransferase; PYRIDOXAL PHOSPHATE.  
FT BINDING 267 267 PYRIDOXAL PHOSPHATE (POTENTIAL).  
SQ SEQUENCE 421 AA; 46166 MW; A4A2E21596E1E16C CRC64;  
  
Query Match 61.1%; Score 33; DB 1; Length 421;  
Best Local Similarity 58.3%; Pred. No. 32;  
Matches 7; Conservative 2; Mismatches 1; Indels 2; Gaps 1;  
  
QY 1 EEVV--PXGMDY 10  
| | | | | | | | | |  
Db 91 EEVILKPRGLDY 102  
  
RESULT 17  
ECB1\_HALEL STANDARD; PRT; 423 AA.  
ID ECB1\_HALEL STANDARD; PRT; 423 AA.  
AC Q9ZEU7;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Diaminobutyrate--pyruvate aminotransferase (EC 2.6.1.46) (L-  
diaminobutyric acid transaminase) (Diaminobutyrate transaminase).  
GN ECB1.  
OS Halomonas elongata.  
OC Bacteria; Proteobacteria; gamma subdivision; Halomonadaceae;  
OC Halomonas.  
OX NCBI\_TaxID=2746;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=DSM 3043;  
RX MEDLINE=99123891; PubMed=9924816;  
RA Canovas D., Vargas C., Calderon M.I., Ventosa A., Nieto J.J.;  
RT "Characterization of the genes for the biosynthesis of the compatible  
solute ectoine in the moderately halophilic bacterium Halomonas  
elongata DSM 3043.";  
RL Syst. Appl. Microbiol. 21:487-497(1998).  
CC -!- CATALYTIC ACTIVITY: L-2,4-diaminobutanate + pyruvate = L-  
aspartate + 4-semialdehyde + L-alanine.  
CC -!- COFACTOR: PYRIDOXAL PHOSPHATE.  
CC -!- PATHWAY: FIRST STEP IN THE BIOSYNTHESIS OF ECTOINE (1,4,5,6-  
TETRAHYDRO-2-METHYL-4-PYRIMIDINE CARBOXYLIC ACID).  
CC -!- SUBUNIT: HOMODIMER (BY SIMILARITY).  
CC -!- SIMILARITY: BELONGS TO CLASS-III OF PYRIDOXAL-PHOSPHATE-DEPENDENT  
CC AMINOTRANSFERASES.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC EMBL: AJ011103; CAA09484.1; -  
CC HSP: P12995; 1QJ3.  
DR

DR Pfam: PF01431; Peptidase\_M13; 1.  
 DR PRINTS; PRO00786; NEPRILYSIN.  
 DR PROSITE; PS00142; ZINC\_PROTEASE; 1.  
 KW Hydrolase; Metalloprotease; Zinc; Glycoprotein; Transmembrane;  
 KW Signal-anchor; Alternative splicing.  
 FT DOMAIN 1 82 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 83 103 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN) (POTENTIAL).  
 FT DOMAIN 104 787 EXTRACELLULAR (POTENTIAL).  
 FT METAL 624 624 ZINC (CATALYTIC) (BY SIMILARITY).  
 FT ACT\_SITE 625 625 BY SIMILARITY.  
 FT METAL 628 628 ZINC (CATALYTIC) (BY SIMILARITY).  
 FT METAL 684 684 ZINC (CATALYTIC) (BY SIMILARITY).  
 FT ACT\_SITE 688 688 PROTON DONOR (BY SIMILARITY).  
 FT CARBOHYD 183 183 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 187 187 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 228 228 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 288 288 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 333 333 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 400 400 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 556 556 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 649 649 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 657 657 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT VARSPLIC 1 63 MOARHAPVQLRWETMDVRKLDPPSASFDVLEKGTLDALL  
 AGERDPWTSSSEGVHTVDQVLS -> MNVALQELGAGSNV  
 EYKRTLREDEAPETPVEGASPDAM (IN ISOFORM ECE-2b).  
 FT VARSPLIC 1 64 MOARHAPVQLRWETMDVRKLDPPSASFDVLEKGTLDALL  
 AGERDPWTSSSEGVHTVDQVLS -> MNVALQELGAGSN (IN ISOFORM ECE-2C).  
 SQ SEQUENCE 787 AA; 89221 MW; CC2D2B0F0EBF7239 CRC64;

Query Match 61.1%; Score 33; DB 1; Length 787;  
 Best Local Similarity 70.0%; Pred. No. 63;  
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 EEVVPXGMDY 10  
 ||| ||||  
 Db 382 EPVVVYGMDY 391

## RESULT 19

FGF3\_MOUSE  
 ID FGF3\_MOUSE STANDARD; PRT; 801 AA.  
 AC Q61851; Q63834; Q61564;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Fibroblast growth factor receptor 3 precursor (EC 2.7.1.112) (FGFR-3)  
 DE (Heparin-binding growth factor receptor).  
 GN FGF3 OR MFR3 OR SAM3.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RX MEDLINE=92355591; PubMed=1379594;  
 RA Ornitz D.M., Leder P.;  
 RT "Ligand specificity and heparin dependence of fibroblast growth  
 RT factor receptors 1 and 3";  
 RL J. Biol. Chem. 267:16305-16311(1992).  
 RN [2]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RC TISSUE=Brain;  
 RX MEDLINE=93177694; PubMed=8382556;  
 RA Kato O., Hattori Y., Sasaki H., Sakamoto H., Fujimoto K.,  
 RA Fujii T., Sugimura T., Terada M.;  
 RT "Isolation of the complementary DNA encoding a mouse heparin-binding  
 RT growth factor receptor with the use of a unique kinase insert  
 RT sequence";  
 RL Cancer Res. 53:1136-1141(1993).  
 RN [3]

RP SEQUENCE OF 242-364 FROM N.A. (ISOFORM 2).  
 RX MEDLINE=94209351; PubMed=7512569;  
 RA Chelliah A.T., McEwen D.G., Werner S., Xu J., Ornitz D.M.;  
 RT "Fibroblast growth factor receptor (FGFR) 3. Alternative splicing in  
 RT immunoglobulin-like domain III creates a receptor highly specific for  
 RT acidic FGF/FGF-1";  
 RL J. Biol. Chem. 269:11620-11627(1994).  
 CC -1- FUNCTION: RECEPTOR FOR ACIDIC AND BASIC FIBROBLAST GROWTH  
 CC FACTORS. PREFERENTIALLY BINDS FGFL.  
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein  
 CC tyrosine phosphate.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- ALTERNATIVE PRODUCTS: 2 isoforms: 1/IIic (shown here) and 2/IIib;  
 CC are produced by alternative splicing.  
 CC -1- TISSUE SPECIFICITY: IN EMBRYO, EXPRESSED IN HEART, LUNG, KIDNEY,  
 CC SKIN, HEAD AND LIVER BUT NOT IN MUSCLE. IN ADULT, HIGHEST LEVELS  
 CC IN BRAIN. ALSO EXPRESSED IN LIVER, LUNG, KIDNEY, TESTIS, OVARY  
 CC AND UTERUS. VERY LOW LEVELS IN HEART, THYMUS, SPLEEN AND MUSCLE.  
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED IN EMBRYO FROM MID-GESTATION AND  
 CC IN ADULT.  
 CC -1- SIMILARITY: BELONGS TO THE FIBROBLAST GROWTH FACTOR RECEPTOR  
 CC FAMILY.  
 CC -1- SIMILARITY: CONTAINS 3 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; M81342; AAA39535.1; -;  
 CC EMBL; S56291; AAB25535.1; -;  
 CC EMBL; L26492; AAA21490.2; -;  
 CC HSSP; P11362; IFGK.  
 CC MGD; MGI:95524; Gfgr3.  
 CC InterPro; IPR000719; Euk\_pkinase.  
 CC InterPro; IPR003006; Ig\_MHC.  
 CC InterPro; IPR003598; Ig\_c2.  
 CC InterPro; IPR001245; Tyr\_pkinase.  
 CC Pfam; PF00047; Ig; 4.  
 CC Pfam; PF00069; pkinase; 1.  
 CC PRINTS; PR00109; TYRKINASE.  
 CC ProDom; PD000001; Euk\_pkinase; 1.  
 CC SMART; SM00408; IGC2; 3.  
 CC SMART; SM00219; TyrKC; 1.  
 CC PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 CC PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
 CC PROSITE; PS00011; PROTEIN\_KINASE\_DOM; 1.  
 CC Receptor; Glycoprotein; Tyrosine-protein kinase; ATP-binding;  
 KW Repeat; Phosphorylation; Transmembrane; Immunoglobulin domain;  
 KW Signal; Alternative splicing.  
 FT SIGNAL 1 20 POTENTIAL.  
 FT CHAIN 21 801 FIBROBLAST GROWTH FACTOR RECEPTOR 3.  
 FT DOMAIN 21 369 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 370 390 POTENTIAL.  
 FT DOMAIN 391 801 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 52 114 IG-LIKE C2-TYPE DOMAIN 1.  
 FT DOMAIN 163 229 IG-LIKE C2-TYPE DOMAIN 2.  
 FT DOMAIN 262 340 IG-LIKE C2-TYPE DOMAIN 3.  
 FT DOMAIN 466 756 PROTEIN KINASE.  
 FT NP\_BIND 472 481 ATP (BY SIMILARITY).  
 FT BINDING 502 502 ATP (BY SIMILARITY).  
 FT ACT\_SITE 611 611 BY SIMILARITY.  
 FT MOD\_RES 642 642 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
 FT DISULFID 59 107 POTENTIAL.  
 FT DISULFID 170 222 POTENTIAL.  
 FT DISULFID 269 333 POTENTIAL.  
 FT CARBOHYD 96 96 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 219 219 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 256 256 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 288 288 N-LINKED (GLCNAC. . .) (POTENTIAL).



FT CARBOHYD 309 309 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 322 322 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT VARSPLIC 305 352 TAGANTTDKEVLVSLHNTFFDAGVTCLAGNSIGFSHHS  
FT ANLVLP -> SWISSEVADARLRANVSOGGEVLCRA  
FT TNEIGVAKAFNLRVHGPOA (IN ISOFORM 2).  
FT P -> L (IN REF. 2).  
FT MISSING (IN REF. 2).  
FT CONFLICT 684 684  
FT CONFLICT 687 687  
SQ SEQUENCE 801 AA; 87758 MW; 68BC110212691705 CRC64;  
  
Query Match 61.1%; Score 33; DB 1; Length 801;  
Best Local Similarity 85.7%; Pred. No. 64;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 5 PXGMDYS 11  
| | | | |  
Db 566 PGMDYS 572  
  
RESULT 20  
CEK2\_CHICK  
ID CEK2\_CHICK STANDARD; PRT; 806 AA.  
AC P18460;  
DT 01-NOV-1990 (Rel. 16, Created)  
DT 01-NOV-1990 (Rel. 16, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Tyrosine kinase receptor CEK2 precursor (EC 2.7.1.112).  
CN CEK2.  
OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianidae;  
OC Gallus.  
OX NCBI\_TaxID=9031;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90332672; PubMed=2165604;  
RA Pasquale E.B.;  
RT "A distinctive family of embryonic protein-tyrosine kinase  
receptors";  
RL Proc. Natl. Acad. Sci. U.S.A. 87:5812-5816(1990).  
RC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein  
tyrosine phosphate.  
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -1- SIMILARITY: BELONGS TO THE FIBROBLAST GROWTH FACTOR RECEPTOR  
FAMILY.  
CC -1- SIMILARITY: CONTAINS 3 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; M35195; AAA48664.1; .  
DR PIR; A35963; A35963.  
DR HSP; P11362; LFCK.  
DR InterPro; IPR000719; Euk\_pkinase.  
DR InterPro; IPR003006; Ig\_MHC.  
DR InterPro; IPR003598; Ig\_c2.  
DR InterPro; IPR001245; Tyr\_pkinase.  
DR Pfam; PF00047; ig; 3.  
DR Pfam; PF00069; pkinase; 1.  
DR PRINTS; PR00109; TYRKINASE.  
DR ProDom; PD000001; Euk\_pkinase; 1.  
DR SMART; SM00408; IGC2; 3.  
DR SMART; SM00219; TYRK; 1.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
DR PROSITE; PS00111; PROTEIN\_KINASE\_DOM; 1.  
KW Receptor; Glycoprotein; Tyrosine-protein kinase; ATP-binding;  
KW Transferase; Phosphorylation; Transmembrane; Immunoglobulin domain;  
KW Repeat; Signal.

FT SIGNAL 1 19  
FT CHAIN 20 806  
FT DOMAIN 364 364  
FT TRANSMEM 365 389  
FT DOMAIN 390 806  
FT DOMAIN 54 114  
FT DOMAIN 163 229  
FT DOMAIN 262 340  
FT DOMAIN 131 141  
FT DOMAIN 466 755  
FT NP\_BIND 472 480  
FT BINDING 502 502  
FT ACT\_SITE 611 611  
FT MOD\_RES 642 642  
FT DISULFID 61 107  
FT DISULFID 170 222  
FT DISULFID 269 333  
FT CARBOHYD 83 83  
FT CARBOHYD 96 96  
FT CARBOHYD 118 118  
FT CARBOHYD 219 219  
FT CARBOHYD 256 256  
FT CARBOHYD 288 288  
FT CARBOHYD 309 309  
FT CARBOHYD 322 322  
SQ SEQUENCE 806 AA; 89730 MW; B38B3C6D5F2314B6 CRC64;  
  
Query Match 61.1%; Score 33; DB 1; Length 806;  
Best Local Similarity 85.7%; Pred. No. 64;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 5 PXGMDYS 11  
| | | | |  
Db 566 PGMDYS 572  
  
RESULT 21  
SULH\_SCHPO  
ID SULH\_SCHPO STANDARD; PRT; 877 AA.  
AC 074377;  
DT 15-JUL-1999 (Rel. 38, Created)  
DT 15-JUL-1999 (Rel. 38, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Probable sulfate permease C3H7.02.  
GN SPBC3H7.02.  
OS Schizosaccharomyces pombe (Fission yeast).  
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
OX NCBI\_TaxID=4896;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=972;  
RX MEDLINE=21848401; PubMed=11859360;  
RA Wood V., Gwilliam R., Rayndream M.A., Lyne M., Lyne R., Stewart A.,  
RA Brooks J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,  
RA Collins M., Brown D., Brown S., Chillingworth T., Churcher C.M.,  
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,  
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,  
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,  
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,  
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,  
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,  
RA Rutherford K., Rutter S., Saunders R., Squares S., Stevens K.,  
RA Skelton J., Simmonds M., Squares R., Walsh S.V., Warren T., Whitehead S.,  
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
RA Woodward J., Volkart G., Aert R., Robben J., Gymonprez B.,  
RA Weijtens J., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,  
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,  
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,  
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,  
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,  
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,

RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,  
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
 RA Certutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,  
 RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.,  
 RT "The genome sequence of Schizosaccharomyces pombe.";  
 RL Nature 415:871-880(2002).  
 CC -!- FUNCTION: HIGH AFFINITY UPTAKE OF SULFATE INTO THE CELL (BY  
 CC SIMILARITY).  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
 CC -!- SIMILARITY: BELONGS TO THE SLC26A FAMILY OF TRANSPORTERS.  
 CC -!- SIMILARITY: CONTAINS 1 STAS DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL: AL031261; CAA20298.1; -.  
 DR InterPro: IPR002845; STAS.  
 DR Pfam: PF00916; Sulfate\_transp.  
 DR Pfam: PF01740; STAS; 1.  
 DR TIGRFAMs: TIGR00815; sulp; 1.  
 DR PROSITE: PS01130; SLC26A; 1.  
 DR PROSITE: PS50801; STAS; 1.  
 KW Transport; Transmembrane.  
 FT TRANSMEM 133 153 POTENTIAL.  
 FT TRANSMEM 161 181 POTENTIAL.  
 FT TRANSMEM 186 206 POTENTIAL.  
 FT TRANSMEM 221 241 POTENTIAL.  
 FT TRANSMEM 243 263 POTENTIAL.  
 FT TRANSMEM 292 312 POTENTIAL.  
 FT TRANSMEM 329 349 POTENTIAL.  
 FT TRANSMEM 384 404 POTENTIAL.  
 FT TRANSMEM 424 444 POTENTIAL.  
 FT TRANSMEM 461 481 POTENTIAL.  
 FT TRANSMEM 484 504 POTENTIAL.  
 FT TRANSMEM 518 538 POTENTIAL.  
 FT TRANSMEM 543 563 POTENTIAL.  
 FT DOMAIN 594 747 STAS.  
 SQ SEQUENCE 877 AA; 96373 MW; 56995A8493371E43 CRC64;

Query Match 61.1%; Score 33; DB 1; Length 877;  
 Best Local Similarity 66.7%; Pred. No. 70;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 QY 3 VVPXGMDYS 11  
 Db 148 VVPQGMSSYA 156  
 ||| | | | | |  
 [1]

RESULT 22  
 ENV\_SFV3L  
 ID ENV\_SFV3L STANDARD; PRT; 982 AA.  
 AC P27399;  
 DT 01-AUG-1992 (Rel. 23, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE ENV polyprotein (Coat polypeptide).  
 GN ENV.  
 OS Simian foamy virus (type 3 / strain LK3) (SFV-3).  
 OC Viruses; Retroviral viruses; Retroviridae; Spumavirus.  
 OX NCBI\_TaxID=11644;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92124734; PubMed=1310187;  
 RA Renne R., Friedl E., Schweizer M., Fleps U., Turek R.,  
 RA Neumann-Haefelin D.;  
 RT "Genomic organization and expression of simian foamy virus type 3

(SFV-3).";  
 RL Virology 186:597-608(1992).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL: M74895; AAA47798.1; ALT\_INIT.  
 DR PIR: C40820; VCLJLK.  
 DR InterPro: IPR005070; Foamy\_env.  
 DR Pfam: PF03408; Foamy\_virus\_ENV; 1.  
 KW Coat protein; Transmembrane; Polyprotein; Glycoprotein.  
 FT TRANSMEM 68 88 I (POTENTIAL).  
 FT TRANSMEM 955 975 II (POTENTIAL).  
 FT CARBOHYD 17 17 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 25 25 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 109 109 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 141 141 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 183 183 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 309 309 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 346 346 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 390 390 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 404 404 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 412 412 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 484 484 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 522 522 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 551 551 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 776 776 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 802 802 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 827 827 N-LINKED (GLCNAC. .) (POTENTIAL).  
 SQ SEQUENCE 982 AA; 113313 MW; 721F2F8929D604FF CRC64;

Query Match 61.1%; Score 33; DB 1; Length 982;  
 Best Local Similarity 66.7%; Pred. No. 79;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 EEVVPXGMD 9  
 Db 44 EEVPIPRMD 52  
 ||| | | | | |  
 [1]

RESULT 23  
 RPOC\_VIBCH STANDARD; PRT; 1401 AA.  
 ID RPOC\_VIBCH STANDARD; PRT; 1401 AA.  
 AC Q9KV29;  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE DNA-directed RNA polymerase beta' chain (EC 2.7.7.6) (Transcriptase  
 DE beta' chain) (RNA polymerase beta' subunit).  
 GN RPOC OR VC0329.  
 OS Vibrio cholerae.  
 OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.  
 OX NCBI\_TaxID=666;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=El Tor N16961 / Serotype O1;  
 RX MEDLINE=20406833; PubMed=10952301;  
 RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,  
 RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,  
 RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,  
 RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,  
 RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,  
 RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,  
 RA Fraser C.M.;  
 RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio  
 RT cholerae.";  
 RL Nature 406:477-483(2000).  
 CC -!- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION

OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS  
SUBSTRATES (BY SIMILARITY).  
-1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +  
[RNA](N).  
-1- SUBUNIT: THE ENZYME CONSISTS OF THE SIGMA CHAIN AND THE CORE  
ENZYME WHICH IS COMPOSED OF 2 ALPHA CHAINS, 1 BETA CHAIN, AND 1  
BETA' CHAIN (BY SIMILARITY).  
-1- SIMILARITY: BELONGS TO THE RNA POLYMERASE BETA' CHAIN FAMILY.  
-----  
THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
EMBL: AE004121; AAF93502.1; --  
HSSP: Q9KRW6; 1HQM.  
TIGR: VC0329; --  
InterPro: IPR000722; RNA\_pol\_A.  
DR InterPro: IPR002879; RNA\_pol\_A2.  
DR Pfam: PF00623; RNA\_pol\_A; 1.  
DR Pfam: PF01854; RNA\_pol\_A2; 2.  
KW Transferase; DNA-directed RNA polymerase; Transcription;  
KW Complete proteome.  
SQ SEQUENCE 1401 AA; 155021 MW; DFDB0F2B5514504F CRC64;  
-----  
Query Match 61.1%; Score 33; DB 1; Length 1401;  
Best Local Similarity 50.0%; Pred. No. 1.2e+02;  
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
-----  
QY 2 EVVPGXGMDYS 11  
Db 581 QIVPKGLPYS 590  
-----  
RESULT 24  
ZEP1\_HUMAN STANDARD; PRT; 2717 AA.  
AC P15822;  
DT 01-APR-1990 (Rel. 14, Last sequence update)  
DT 01-APR-1990 (Rel. 14, Last annotation update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Zinc finger protein 40 (Human immunodeficiency virus type I enhancer-  
binding protein 1) (HIV-EPI) (Major histocompatibility complex binding  
protein 1) (MBP-1) (Positive regulatory domain II binding factor 1)  
(PRDII-BF1).  
DE (PRDII-BF1).  
GN HIVP1 OR ZNF40.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90169514; PubMed=2106471;  
RA Fan C.M., Maniatis T.;  
RT "A DNA-binding protein containing two widely separated zinc finger  
motifs that recognize the same DNA sequence."  
RL Genes Dev. 4:29-42(1990).  
RN [2]  
RP STRUCTURE BY NMR OF 2113-2142.  
RX MEDLINE=91064333; PubMed=2248949;  
RA Omichinski J.G., Clore G.M., Appella E., Sakaguchi K.,  
RA Gronenborn A.M.;  
RT "High-resolution three-dimensional structure of a single zinc finger  
from a human enhancer binding protein in solution."  
RL Biochemistry 29:9324-9334(1990).  
RN [3]  
RP STRUCTURE BY NMR OF 2087-2142.  
RX MEDLINE=92232684; PubMed=1567844;  
RA Omichinski J.G., Clore G.M., Robien M., Sakaguchi K., Appella E.,  
RA Gronenborn A.M.;

High-resolution solution structure of the double Cys2His2 zinc  
finger from the human enhancer binding protein MBP-1.";  
Biochemistry 31:3907-3917(1992).  
-1- FUNCTION: THIS PROTEIN SPECIFICALLY BINDS TO THE DNA SEQUENCE  
5'-GGGACTTCC-3' WHICH IS FOUND IN THE ENHANCER ELEMENTS OF  
NUMEROUS VIRAL PROMOTERS SUCH AS THOSE OF SV40, CMV, OR HIV1.  
IN ADDITION, RELATED SEQUENCES ARE FOUND IN THE ENHANCER ELEMENTS  
OF A NUMBER OF CELLULAR PROMOTERS, INCLUDING THOSE OF THE CLASS I  
MHC, INTERLEUKIN-2 RECEPTOR, AND INTERFERON-BETA GENES. IT MAY ACT  
IN T-CELL ACTIVATION.  
-1- SUBCELLULAR LOCATION: Nuclear.  
-1- INDUCTION: BY MITOGEN AND PHORBOL ESTER.  
-1- DOMAIN: CONTAINS TWO SETS OF 2 ZINC-FINGERS, WHICH ARE WIDELY  
SEPARATED AND RECOGNIZE THE SAME DNA SEQUENCE. THERE IS A FIFTH  
ZINC-FINGER IN-BETWEEN.  
-1- SIMILARITY: STRONG, TO HIVP2.  
-----  
THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
EMBL: X51435; CAA35798.1; --  
PIR: A34203; A34203.  
PDB: 3ZNE; 15-JAN-92.  
PDB: 4ZNF; 15-JAN-92.  
PDB: 1BBO; 31-OCT-93.  
DR TRANSFAC: T00497; --  
DR Genew; HGNC:4920; HIVP1.  
DR MIM: 194540; --  
DR InterPro: IPR000822; Znf\_C2H2.  
DR Pfam: PF00096; zf-C2H2; 5.  
DR PRINTS: PR00048; ZINCFINGER.  
DR SMART: SM00355; Znf\_C2H2; 4.  
DR PROSITE: PS00028; ZINC\_FINGER\_C2H2\_1; 4.  
DR PROSITE: PS00157; ZINC\_FINGER\_C2H2\_2; 4.  
KW Transcription regulation; Zinc-finger; Metal-binding; DNA-binding;  
KW Nuclear protein; Repeat; 3D-structure.  
KW DOMAIN 406 456 ZINC FINGERS.  
FT ZN\_FING 406 428 C2H2-TYPE.  
FT ZN\_FING 434 456 C2H2-TYPE.  
FT DOMAIN 806 806 POLY-SER.  
FT ZN\_FING 958 981 C2HC-TYPE (POTENTIAL).  
FT DOMAIN 2087 2139 ZINC FINGERS.  
FT ZN\_FING 2087 2109 C2H2-TYPE.  
FT ZN\_FING 2115 2139 C2H2-TYPE.  
FT STRAND 2088 2088  
FT TURN 2090 2092  
FT STRAND 2095 2095  
FT HELIX 2099 2108  
FT TURN 2109 2109  
FT STRAND 2115 2116  
FT STRAND 2123 2124  
FT HELIX 2127 2135  
SQ SEQUENCE 2717 AA; 297217 MW; D45D3CA951FEA561 CRC64;  
-----  
Query Match 61.1%; Score 33; DB 1; Length 2717;  
Best Local Similarity 66.7%; Pred. No. 2.3e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
-----  
QY 3 VVPXGMDYS 11  
Db 2405 VVPAGLTYS 2413  
-----  
RESULT 25  
ET2A\_XENLA STANDARD; PRT; 472 AA.  
ID ET2A\_XENLA  
AC P19102;  
DT 01-NOV-1990 (Rel. 16, Created)

DT 01-OCT-1993 (Rel. 27, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE C-ETS-2A protein.  
GN ETS2A OR ETS-2A.  
OS Xenopus laevis (African clawed frog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;  
OC Xenopodinae; Xenopus.  
OX NCBI\_TaxID=8355;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92158632; PubMed=1741266;  
RA Burdett L.A., Qi S.M., Chen Z.Q., Lautenberger J.A., Papas T.S.;  
RT "Characterization of the cDNA sequences of two Xenopus ets-2 proto-  
oncogenes.";  
RL Nucleic Acids Res. 20:371-371(1992).  
RN [2]  
RP SEQUENCE OF 121-472 FROM N.A.  
RC TISSUE=Oocyte;  
RX MEDLINE=90356411; PubMed=2201951;  
RA Wolff C.M., Stiegler P., Baltzinger M., Meyer D., Ghysdael J.,  
RA Stehelin D., Befort N., Remy P.;  
RT "Isolation of two different c-ets-2 proto-oncogenes in Xenopus  
laevis.";  
RL Nucleic Acids Res. 18:4603-4604(1990).  
RN [3]  
RP SEQUENCE OF 121-472 FROM N.A.  
RX MEDLINE=92088972; PubMed=1751411;  
RA Wolff C.M., Stiegler P., Baltzinger M., Meyer D., Ghysdael J.,  
RA Stehelin D., Befort N., Remy P.;  
RT "Cloning, sequencing, and expression of two Xenopus laevis c-ets-2  
proto-oncogenes.";  
RL Cell Growth Differ. 2:447-456(1991).  
CC -!- SUBCELLULAR LOCATION: Nuclear.  
CC -!- SIMILARITY: BELONGS TO THE ETS FAMILY.  
CC -!- SIMILARITY: CONTAINS 1 POINTED (PNT) DOMAIN.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; M81683; AAA49705.1; -;  
DR EMBL; X51826; CAA36124.1; -;  
DR PIR; S10994; S10994.  
DR PIR; S28824; S28824.  
DR HSP; P14921; 2STT.  
DR TRANSPAC; T02041; -;  
DR InterPro; IPR000418; Ets.  
DR InterPro; IPR002341; HSF\_ETS.  
DR InterPro; IPR003118; SAM\_PNT.  
DR Pfam; PF00178; Ets; 1.  
DR Pfam; PF02198; SAM\_PNT; 1.  
DR PRINTS; PR00454; ETSDOMAIN.  
DR SMART; SM00413; ETS; 1.  
DR SMART; SM00251; SAM\_PNT; 1.  
DR PROSITE; PS00345; ETS\_DOMAIN\_1; 1.  
DR PROSITE; PS00346; ETS\_DOMAIN\_2; 1.  
DR PROSITE; PS00061; ETS\_DOMAIN\_3; 1.  
KW DNA-binding; Nuclear protein.  
FT DOMAIN 87 170 POINTED.  
FT DNA\_BIND 366 446 ETS-DOMAIN.  
SQ SEQUENCE 472 AA; 53894 MW; E0E808B5E6BF111 CRC64;  
Query Match 60.2%; Score 32.5; DB 1; Length 472;  
Best Local Similarity 58.3%; Pred. No. 46;  
Matches 7; Conservative 2; Mismatches 2; Indels 1; Gaps 1;  
OY 1 EEVVPXGMD-YS 11  
|:|:|:|:|

Db 49 EQAVPTGLDSYS 60

Search completed: June 10, 2003, 13:40:19  
Job time : 5.5 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2003, 13:30:25 ; Search time 25.7857 Seconds  
(without alignments)  
87.898 Million cell updates/sec

Title: US-09-909-164-8  
Perfect score: 54  
Sequence: 1 EVVFXGMDYS 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues  
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 75 summaries

Database : SPTREMBL21.\*

- 1: sp\_archaea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mhc.\*
- 8: sp\_organelle.\*
- 9: sp\_phase.\*
- 10: sp\_plant.\*
- 11: sp\_rodent.\*
- 12: sp\_virus.\*
- 13: sp\_vertebrate.\*
- 14: sp\_unclassified.\*
- 15: sp\_rvirus.\*
- 16: sp\_bacteriap.\*
- 17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	40	74.1	156	3 Q12479	Q12479 saccharomyc
2	38	70.4	363	17 Q30260	Q30260 archaeoglob
3	38	70.4	1063	16 Q8RG86	Q8RG86 fusobacteri
4	36	66.7	341	10 Q22081	Q22081 citrus unsh
5	36	66.7	348	10 Q22096	Q22096 citrus unsh
6	36	66.7	452	10 Q8W568	Q8W568 arabidopsis
7	36	66.7	460	10 Q9C9T7	Q9C9T7 arabidopsis
8	36	66.7	1047	10 P93782	P93782 saccharum o
9	36	66.7	1083	10 Q9SN30	Q9SN30 arabidopsis
10	36	66.7	1084	10 Q43010	Q43010 oryza sativ
11	36	66.7	1100	10 Q8S064	Q8S064 oryza sativ
12	35	64.8	219	5 Q9GQ84	Q9GQ84 eriocheir s
13	35	64.8	253	16 Q8XPA8	Q8XPA8 clostridium
14	35	64.8	298	2 Q52367	Q52367 rhizobium t
15	35	64.8	425	5 Q9XVK4	Q9XVK4 caenorhabdi
16	35	64.8	433	16 Q9A382	Q9A382 caulobacter

17	35	64.8	440	17 Q9YFI3	Q9YFI3 aeropyrum p
18	35	64.8	511	2 Q52680	Q52680 escherichia
19	35	64.8	517	16 Q8XZL5	Q8XZL5 ralstonia s
20	35	64.8	745	5 Q95P46	Q95P46 carcinus ma
21	35	64.8	1031	5 Q9U6A3	Q9U6A3 callinectes
22	35	64.8	1150	5 Q17704	Q17704 caenorhabdi
23	35	64.8	1410	2 Q52673	Q52673 escherichia
24	35	64.8	1420	2 Q52666	Q52666 escherichia
25	35	64.8	1474	17 Q27146	Q27146 methanobact
26	35	64.8	1828	16 Q98K29	Q98K29 rhizobium l
27	35	64.8	2778	5 Q9V9T6	Q9V9T6 drosophila
28	34.5	63.9	748	4 Q8TBJ7	Q8TBJ7 homo sapien
29	34	63.0	143	5 Q9VSY8	Q9VSY8 drosophila
30	34	63.0	154	10 Q9SBB8	Q9SBB8 cryza sativ
31	34	63.0	215	16 Q8RSL5	Q8RSL5 thermoanaer
32	34	63.0	290	16 Q8U7J0	Q8U7J0 agrobacteri
33	34	63.0	296	17 Q9YET8	Q9YET8 aeropyrum p
34	34	63.0	357	17 Q29920	Q29920 archaeoglob
35	34	63.0	366	17 Q29451	Q29451 archaeoglob
36	34	63.0	387	16 Q98FX1	Q98FX1 rhizobium l
37	34	63.0	543	3 Q8TPF4	Q8TPF4 trichoderma
38	34	63.0	558	16 Q8R822	Q8R822 thermoanaer
39	34	63.0	565	16 Q9CIN1	Q9CIN1 lactococcus
40	34	63.0	587	16 Q9JZP8	Q9JZP8 neisseria m
41	34	63.0	906	10 Q9ZSY4	Q9ZSY4 arabidopsis
42	34	63.0	908	10 Q9FJK8	Q9FJK8 arabidopsis
43	34	63.0	908	10 Q8W4J9	Q8W4J9 arabidopsis
44	34	63.0	908	10 Q9ZSY3	Q9ZSY3 arabidopsis
45	34	63.0	909	10 Q9M5A1	Q9M5A1 arabidopsis
46	34	63.0	1062	2 P95422	P95422 pseudomonas
47	34	63.0	1062	16 Q910Y8	Q910Y8 pseudomonas
48	34	63.0	1062	1 Q74056	Q74056 cenarchaeum
49	33	61.1	78	6 Q9XST4	Q9XST4 canis famil
50	33	61.1	97	17 Q97VR9	Q97VR9 sulfolobus
51	33	61.1	128	17 Q97US8	Q97US8 sulfolobus
52	33	61.1	172	13 Q02528	Q02528 oryzias lat
53	33	61.1	175	10 Q8VY88	Q8VY88 arabidopsis
54	33	61.1	183	2 Q9S110	Q9S110 sweet potat
55	33	61.1	184	17 Q97XJ2	Q97XJ2 sulfolobus
56	33	61.1	193	5 Q8SWL4	Q8SWL4 encephalito
57	33	61.1	209	13 Q9DEL6	Q9DEL6 brachydanio
58	33	61.1	210	10 Q65890	Q65890 cyclotella
59	33	61.1	217	4 Q00404	Q00404 homo sapien
60	33	61.1	225	10 Q40129	Q40129 lycopersico
61	33	61.1	230	17 Q9P9L6	Q9P9L6 pyrobaculum
62	33	61.1	247	16 Q92U66	Q92U66 rhizobium m
63	33	61.1	257	10 Q9C6J0	Q9C6J0 arabidopsis
64	33	61.1	262	17 Q97YB8	Q97YB8 sulfolobus
65	33	61.1	267	17 Q97Y57	Q97Y57 sulfolobus
66	33	61.1	299	4 Q9UEF9	Q9UEF9 homo sapien
67	33	61.1	299	17 Q97TV4	Q97TV4 sulfolobus
68	33	61.1	299	17 Q9DWN4	Q9DWN4 sulfolobus
69	33	61.1	307	17 Q9HP60	Q9HP60 halobacteri
70	33	61.1	396	17 Q8TIT9	Q8TIT9 methanosarc
71	33	61.1	441	17 Q9HIW6	Q9HIW6 thermoplasm
72	33	61.1	556	4 Q43733	Q43733 homo sapien
73	33	61.1	577	15 Q9QBR8	Q9QBR8 simian foam
74	33	61.1	577	15 Q9QBR7	Q9QBR7 simian foam
75	33	61.1	577	15 Q9QBR6	Q9QBR6 simian foam

## ALIGNMENTS

RESULT 1  
Q12479  
ID Q12479  
AC Q12479  
DT 01-NOV-1996 (TREMREL. 01, Created)  
DT 01-NOV-1996 (TREMREL. 01, Last sequence update)  
DT 01-JUN-2002 (TREMREL. 21, Last annotation update)  
DE ORF YOR013W.  
GN YOR013W.

OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
 OX NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA De haan M., Grivell L.A., Maarse A.C.;  
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA MIPS;  
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA STRAIN-FY1679;  
 RL De haan M., Maarse A.C., Grivell L.A.;  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE-94019318; PubMed=8413243;  
 RX Dumont M.E., Schlichter J.B., Cardillo T.S., Hayes M.K., Bethlenny G.,  
 RA Sherman F.;  
 RT "CYC2 encodes a factor involved in mitochondrial import of yeast  
 RT cytochrome c.";  
 RL Mol. Cell. Biol. 13:6442-6451(1993).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RA STRAIN-FY1679;  
 RX MEDLINE-94169519; PubMed=7764548;  
 RA Lee Y.S., Shimizu J., Yoda K., Yanasaki M.;  
 RT "Molecular cloning of a gene, DHS1, which complements a drug-  
 RT hypersensitive mutation of the yeast Saccharomyces cerevisiae.";  
 RL Biosci. Biotechnol. Biochem. 58:391-395(1994).  
 DR EMBL; 274920; CAA99201.1; -;  
 DR EMBL; X87331; CAA60762.1; -;  
 DR SGD; S0005539; YOR013W.  
 SQ SEQUENCE 156 AA; 17881 MW; 380442B74C272B41 CRC64;  
 Query Match 74.1%; Score 40; DB 3; Length 156;  
 Best Local Similarity 77.8%; Pred. No. 1.7;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 2 EVVFXGMDY 10  
 Db 50 EVVPLGMDY 58  
 RESULT 2  
 O30260  
 ID O30260 PRELIMINARY; PRT; 363 AA.  
 AC O30260;  
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 DE Hypothetical protein AF2411.  
 GN AF2411.  
 OS Archaeoglobus fulgidus.  
 OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;  
 OC Archaeoglobaceae; Archaeoglobus.  
 OX NCBI\_TaxID=2234;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA STRAIN-VC-16 / DSM 4304 / ATCC 49558;  
 RX MEDLINE=98049343; PubMed=9389475;  
 RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,  
 RA Ketchum K.A., Dodson R.J., Winn M., Hickey E.K., Peterson J.D.,  
 RA Richardson D.L., Kralavag A.R., Graham D.E., Kyrpides N.C.,  
 RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,  
 RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,  
 RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,  
 RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,  
 RA Cotton M.D., Spriggs T., Artiach P., Kaine B.P., Sykes S.M.,

RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,  
 RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,  
 RA Venter J.C.;  
 RT "The complete genome sequence of the hyperthermophilic, sulphate-  
 RT reducing archaeon Archaeoglobus fulgidus.";  
 RL Nature 390:364-370(1997).  
 DR EMBL; AE001109; AAB91255.1; -;  
 DR TIGR; AF2411; -;  
 DR InterPro; IPR002103; Bac\_luciferase.  
 DR Pfam; PF00296; bac\_luciferase; 1.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 363 AA; 41736 MW; 0E976EAE788F4803 CRC64;  
 Query Match 70.4%; Score 38; DB 17; Length 363;  
 Best Local Similarity 54.5%; Pred. No. 12;  
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 EEVVPXGMDYS 11  
 Db 120 ENIVPYGIDFS 130  
 RESULT 3  
 Q8RG86  
 ID Q8RG86 PRELIMINARY; PRT; 1063 AA.  
 AC Q8RG86;  
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE Carbamoyl-phosphate synthase large chain (EC 6.3.5.5).  
 GN FN0422.  
 OS Fusobacterium nucleatum (subsp. nucleatum).  
 OC Bacteria; Fusobacteria; Fusobacterium.  
 OX NCBI\_TaxID=76856;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 25586;  
 RX MEDLINE=21886394; PubMed=11889109;  
 RA Kapral V., Anderson I., Ivanova N., Reznik G., Lykidis A.,  
 RA Bhattacharyya A., Bartman A., Gardner W., Grechkin G., Zhu L.,  
 RA Vasieva O., Chu L., Kogan Y., Chaga O., Goltzman E., Bernal A.,  
 RA Larsen N., D'Souza M., Walunas T., Pusch G., Haselkorn R.,  
 RA Fongstein M., Kyrpides N., Overbeek R.;  
 RT "Genome sequence and analysis of the oral bacterium Fusobacterium  
 RT nucleatum strain ATCC 25586.";  
 RL J. Bacteriol. 184:2005-2018(2002).  
 DR EMBL; AE010554; AAL94625.1; -;  
 KW Ligase; Complete proteome.  
 SQ SEQUENCE 1063 AA; 118008 MW; 39700E10B7CCE411 CRC64;  
 Query Match 70.4%; Score 38; DB 16; Length 1063;  
 Best Local Similarity 60.0%; Pred. No. 39;  
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
 QY 2 EVVFXGMDYS 11  
 Db 195 EIVPGLNYS 204  
 RESULT 4  
 O22081  
 ID O22081 PRELIMINARY; PRT; 341 AA.  
 AC O22081;  
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE Sucrose-phosphate synthase (Fragment).  
 GN CITSPS2.  
 OS Citrus unshiu (Satsuma orange).  
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Sapindales; Rutaceae; Citrus.  
 OX NCBI\_TaxID=55188;

```

RN RP SEQUENCE FROM N.A.
RC STRAIN=CV. MIYAGAWA-WASE; TISSUE=JUICE SACS AND SEGMENT EPIDERMIS;
RX MEDLINE=96439842; PubMed=8842155;
RA Komatsu A., Takanokura Y., Omura M., Akihama T.;
RT "Cloning and molecular analysis of cDNA encoding three sucrose
RT phosphate synthase isoforms from a citrus fruit (Citrus unshiu
RT Marc.).";
RL Mol. Gen. Genet. 252:346-351(1996).
[2]
RN RP SEQUENCE FROM N.A.
RC STRAIN=CV. MIYAGAWA-WASE; TISSUE=JUICE SACS AND SEGMENT EPIDERMIS;
RA Komatsu A., Takanokura Y., Moriguchi T., Omura M., Akihama T.;
RT "Differential expression of three sucrose-phosphate synthase isoforms
RT during sucrose accumulation in citrus fruits (Citrus unshiu Marc.).";
RL Plant Sci. 140:169-178(1999).
DR ENBL; AB006319; BAA23215.1;
FT NON_TER 1 341
FT NON_TER 341 341
SQ SEQUENCE 341 AA; 38136 MW; 61417A69C4560777 CRC64;

Query Match 66.7%; Score 36; DB 10; Length 341;
Best Local Similarity 66.7%; Pred. No. 29;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMDYS 11
Db 228 VIPPGMDFS 236

RESULT 5
O22096 PRELIMINARY; PRT; 348 AA.
AC O22096;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Sucrose-phosphate synthase (Fragment).
GN CITSPS3.
OS Citrus unshiu (Satsuma orange).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Sapindales; Rutaceae; Citrus.
OX NCBI_TaxID=55188;
[1]
RN RP SEQUENCE FROM N.A.
RC TISSUE=JUICE SACS AND SEGMENT EPIDERMIS;
RX MEDLINE=96439842; PubMed=8842155;
RA Komatsu A., Takanokura Y., Omura M., Akihama T.;
RT "Cloning and molecular analysis of cDNA encoding three sucrose
RT phosphate synthase isoforms from a citrus fruit (Citrus unshiu
RT Marc.).";
RL Mol. Gen. Genet. 252:346-351(1996).
[2]
RN RP SEQUENCE FROM N.A.
RC TISSUE=JUICE SACS AND SEGMENT EPIDERMIS;
RA Komatsu A., Takanokura Y., Moriguchi T., Omura M., Akihama T.;
RT "Differential expression of three sucrose-phosphate synthase isoforms
RT during sucrose accumulation in citrus fruits (Citrus unshiu Marc.).";
RL Plant Sci. 140:169-178(1999).
DR ENBL; AB006660; BAA23071.1;
FT NON_TER 1 348
FT NON_TER 348 348
SQ SEQUENCE 348 AA; 38556 MW; EE1C21EBA6FF5C5E CRC64;

Query Match 66.7%; Score 36; DB 10; Length 348;
Best Local Similarity 66.7%; Pred. No. 30;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMDYS 11
Db 234 VIPPGMDFS 242

```

```

RESULT 6
Q8W568 PRELIMINARY; PRT; 452 AA.
AC Q8W568;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Atlg73750/F25P22.17.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
[1]
RN RP SEQUENCE FROM N.A.
RA Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,
RA Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
RA Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,
RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,
RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,
RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
RA Ecker J.R.;
RT "Arabidopsis cDNA clones.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR ENBL; AF419606; AAL31938.1;
DR InterPro; IPR002471; Prol_endopep_ser.
DR InterPro; IPR000379; Ser_estrs_site.
DR PROSITE; PS00708; PRO_ENDOPEP_SER; UNKNOWN1.
SQ SEQUENCE 452 AA; 49682 MW; A159955B21742C4A CRC64;

Query Match 66.7%; Score 36; DB 10; Length 452;
Best Local Similarity 70.0%; Pred. No. 40;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMDY 10
Db 210 EEDVPSAMDY 219

RESULT 7
Q9C9T7 PRELIMINARY; PRT; 460 AA.
AC Q9C9T7;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Hypothetical 50.6 kDa protein.
GN F25P22.17.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
[1]
RN RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RX MEDLINE=21016719; PubMed=11130712;
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y.,
RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
RA Dunn P., Etgu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA Gull J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA Lin X., Liu S.X., Liu Z.A., Luyos J.S., Maiti R., Marziani A.,
RA Millitscher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,

```

```

RA Utterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
RT thaliana.";
RL Nature 408:816-820(2000).
DR EMBL; AC012679; AAG52073.1; -.
DR InterPro; IPR002471; ProL_endopep_ser.
DR PROSITE; PS003379; Ser_estrs_site.
DR PROSITE; PS00708; PRO-ENDOPEP_SER; UNKNOWN_1.
KW Hypothetical protein.
SQ SEQUENCE 460 AA; 50564 MW; E94B27B5C4B249EC CRC64;

Query Match 66.7%; Score 36; DB 10; Length 460;
Best Local Similarity 70.0%; Pred. No. 41;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMDY 10
DB 218 EEDVPSANDY 227

RESULT 8
P93782 PRELIMINARY; PRT; 1047 AA.
AC P93782;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Sucrose-phosphate synthase (EC 2.4.1.14) (Fragment).
GN SOSPS1.
OS Saccharum officinarum (Sugarcane).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoideae; Andropogoneae; Saccharum.
OX NCBI_TaxID=4547;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LEAF;
RA Sugiharto B., Sakakibara H., Sugiyama T.;
RT "Differential Expression of Two Genes for Sucrose-Phosphate Synthase
RT in Sugarcane: Molecular Cloning of the cDNAs and Comparative Analysis
RT of Gene Expression.";
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB001337; BAA19241.1; -.
DR InterPro; IPR001296; Glycos_transf_1.
DR Pfam; PF00534; Glycos_transf_1.1.
KW Glycosyltransferase; Transferase.
FT NON_TER
SQ SEQUENCE 1047 AA; 116379 MW; D0EDB34961E1D83D CRC64;

Query Match 66.7%; Score 36; DB 10; Length 1047;
Best Local Similarity 66.7%; Pred. No. 1e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMDYS 11
DB 414 VIPPGMDFS 422

RESULT 9
Q9SN30 PRELIMINARY; PRT; 1083 AA.
AC Q9SN30;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Sucrose-phosphate synthase-like protein (EC 2.4.1.14).
GN F28M11.40 OR AT4G10120.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;

```

```

RN SEQUENCE FROM N.A.
RP Bevan M., Murphy G., Ridley P., Hudson S., Bancroft I., Mewes H.W.,
RA Mayer K.F.X., Lencke K., Schueller C.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL049487; CAB39764.1; -.
DR EMBL; AL161516; CAB78135.1; -.
DR InterPro; IPR001296; Glycos_transf_1.
DR Pfam; PF00534; Glycos_transf_1.1.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 1083 AA; 122688 MW; EFC126FCA2137BB1 CRC64;

Query Match 66.7%; Score 36; DB 10; Length 1083;
Best Local Similarity 66.7%; Pred. No. 1.1e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMDYS 11
DB 483 VIPPGMDFS 491

RESULT 10
Q43010 PRELIMINARY; PRT; 1084 AA.
AC Q43010;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE L-lactate dehydrogenase (EC 1.1.1.27) (LDH).
GN SP51.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. JAPONICA;
RA Sakamoto M., Satozawa T., Kishimoto N., Higo K., Shimada H.,
RA Fujimura T.;
RT "Structure and RFLP mapping of a rice sucrose phosphate synthase (SPS)
RT gene that is specifically expressed in the source organ.";
RL Plant Sci. 112:207-217(1995).
CC -|- CATALYTIC ACTIVITY: (S)-LACTATE + NAD(+) = PYRUVATE + NADH.
CC -|- PATHWAY: ANAEROBIC GLYCOLYSIS; FINAL STEP.
CC -|- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).
DR EMBL; D45890; BAA08304.1; -.
DR InterPro; IPR001296; Glycos_transf_1.
DR InterPro; IPR001557; L_LDH.
DR Pfam; PF00534; Glycos_transf_1.1.
DR PROSITE; PS00064; L_LDH; 1.
KW Glycolysis; NAD; Oxidoreductase.
SQ SEQUENCE 1084 AA; 119417 MW; 4D0D79AFEDD2F92B CRC64;

Query Match 66.7%; Score 36; DB 10; Length 1084;
Best Local Similarity 66.7%; Pred. No. 1.1e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMDYS 11
DB 453 VIPPGMDFS 461

RESULT 11
Q8S064

```



```

ID Q8S064 PRELIMINARY; PRT; 1100 AA.
AC Q8S064;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DE Putative sucrose-phosphate synthase.
GN P0678F11.14.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoidae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
EMBL; AP003437; BAB86107.1; -.
SQ SEQUENCE 1100 AA; 121170 MW; 084F4604BA389CAD CRC64;

Query Match 66.7%; Score 36; DB 10; Length 1100;
Best Local Similarity 66.7%; Pred. No. 1.1e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMDYS 11
|:|:|:|:|
Db 469 VIPPGMDFS 477

RESULT 12
Q9GQ04 PRELIMINARY; PRT; 219 AA.
AC Q9GQ04;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DE Na+/K+/2Cl-cotransporter (Fragment).
OS Eriocheir sinensis (Chinese mitten crab).
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC Eubrachyura; Grapsoidae; Varunidae; Eriocheir.
OX NCBI_TaxID=95602;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=GILL;
RA Wehrhau D., Towle D.W.;
RT "Na+/H+-exchanger and Na+/K+/2Cl- cotransporter are expressed in
RL gills of the euryhaline Chinese crab Eriocheir sinensis.";
RL Comp. Biochem. Physiol. 126:S158-S158(2000).
DR EMBL; AF301160; AAG39938.1; -.
FT NON_TER 1
FT NON_TER 219
SQ SEQUENCE 219 AA; 24159 MW; 599442DA26F0D3F1 CRC64;

Query Match 64.8%; Score 35; DB 5; Length 219;
Best Local Similarity 75.0%; Pred. No. 29;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 VPXGMDYS 11
|:|:|:|:|
Db 107 VPQGLDYS 114

RESULT 13
Q8XPA8 PRELIMINARY; PRT; 253 AA.
AC Q8XPA8;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DE Hypothetical protein CPE0057.

```

```

GN CPE0057.
OS Clostridium perfringens.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
OC Clostridiales; Clostridiaceae; Clostridium.
OX NCBI_TaxID=1502;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=13 / TYPE A;
RX PubMed=11792842;
RA Shimizu T., Ohtani K., Hirakawa H., Ohshima K., Yamashita A.,
RA Shiba T., Ogasawara N., Hattori M., Kuhara S., Hayashi H.;
RT "Complete genome sequence of Clostridium perfringens, an anaerobic
RT flesh-eater.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:996-1001(2002).
DR EMBL; AF003185; BAB79763.1; -.
DR InterPro; IPR000205; NAD_binding.
DR InterPro; IPR000594; ThiF_domain.
DR Pfam; PF00899; ThiF; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 253 AA; 27714 MW; 814DF79D3E0D7486 CRC64;

Query Match 64.8%; Score 35; DB 16; Length 253;
Best Local Similarity 50.0%; Pred. No. 34;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMDY 10
|:|:|:|:|
Db 108 EEIIPDDVDY 117

RESULT 14
O52367 PRELIMINARY; PRT; 298 AA.
AC O52367;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Acyl-alcohol dehydrogenase homolog (Fragment).
GN XYLBI.
OS Rhizobium tropici.
OG Plasmid pRtCFN299a.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=398;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CFN299;
RA Rosenblueth M., Hynes M.F., Martinez-Romero E.;
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
CC -|- COFACTOR: ZINC (BY SIMILARITY).
CC -|- SIMILARITY: BELONGS TO THE ZINC-CONTAINING ALCOHOL DEHYDROGENASE
CC FAMILY.
DR EMBL; AF036920; AAC04779.1; -.
DR HSPSP; P07846; ISDG.
DR InterPro; IPR002328; ADH_zinc.
DR InterPro; IPR002085; Adh_zn_family.
DR InterPro; IPR000205; NAD_binding.
DR Pfam; PF00107; adh_zinc; 1.
DR PROSITE; PS00059; ADH_ZINC; 1.
KW Oxidoreductase; Plasmid; Zinc.
FT NON_TER 298
SQ SEQUENCE 298 AA; 31092 MW; 49B2F8117C33AE87 CRC64;

Query Match 64.8%; Score 35; DB 2; Length 298;
Best Local Similarity 50.0%; Pred. No. 41;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 EVVVPXGMDYS 11
|:|:|:|:|
Db 250 EIIPEGADFS 259

RESULT 15

```

Q9XVK4  
ID Q9XVK4 PRELIMINARY; PRT; 425 AA.  
AC Q9XVK4;  
DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
DE R10D12.10 protein.  
GN R10D12.10  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Pelodierinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Percy C.M.;  
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99069613; PubMed=9851916;  
RA none;  
RT "Genome sequence of the nematode C.elegans: A platform for  
investigating biology."  
RL Science 282:2012-2018(1998).  
DR EMBL; Z81109; CAB03241.1;  
DR InterPro; IPR000719; Euk\_pkinase.  
DR Pfam; PF00069; pkinase; 1.  
DR ProDom; PD000001; Euk\_pkinase; 1.  
DR PROSITE; PS00011; PROTEIN\_KINASE\_DOM; 1.  
KW ATP-binding; Transferase.  
SQ SEQUENCE 425 AA; 49410 MW; 5D96E29B08C8E9D6 CRC64;

Query Match 64.8%; Score 35; DB 5; Length 425;  
Best Local Similarity 50.0%; Pred. No. 61;  
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EEVVPXGMDY 10  
I::I I: I  
Db 335 EQIVPGGLQY 344

RESULT 16  
Q9A382 PRELIMINARY; PRT; 433 AA.  
ID Q9A382;  
AC Q9A382;  
DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
DE Peptidoglycan-binding protein, putative.  
GN CC3322.  
OS Caulobacter crescentus.  
OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;  
OC Caulobacter.  
OX NCBI\_TaxID=155892;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=ATCC 19089 / CBI5;  
MEDLINE=21173698; PubMed=11259647;  
RA Eisen W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,  
Nierman W.C., Newton A., Stephens C., Phadke N.D., Ely B.,  
Pollock I., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,  
DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,  
Kolony J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,  
Utterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,  
Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;  
RT "Complete genome sequence of Caulobacter crescentus."  
Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).  
RL EMBL; AE005994; AAK25284.1;  
DR HSSP; P41052; 1LTM.  
DR TIGR; CC3322;  
DR InterPro; IPR002477; PG\_binding.  
DR Pfam; PF01471; PG\_binding\_1; 1.  
KW Complete proteome.  
SQ SEQUENCE 433 AA; 46169 MW; F7DD18B9F4CA10A7 CRC64;

Query Match 64.8%; Score 35; DB 16; Length 433;  
Best Local Similarity 54.5%; Pred. No. 62;  
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 EEVVPXGMDY 11  
I::I I: I  
Db 266 EVILPPGFDY 276

RESULT 17  
Q9YFI3 PRELIMINARY; PRT; 440 AA.  
ID Q9YFI3  
AC Q9YFI3;  
DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE 440AA long hypothetical alkaline protease.  
GN APE0263.  
OS Aeropyrum pernix.  
OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcales;  
OC Desulfurococaceae; Aeropyrum.  
OX NCBI\_TaxID=56636;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=KJ;  
RC MEDLINE=99310339; PubMed=10382966;  
RA Kwarabavasi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,  
Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankai A., Kosugi H.,  
Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,  
Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudooh Y.,  
Yamazaki J., Kishida N., Oguchi A., Aoki K.-I., Kubota K.,  
Nakamura Y., Nomura N., Sako Y., Kikuchi H.;  
RT "Complete genome sequence of an aerobic hyper-thermophilic  
crenarchaeon, Aeropyrum pernix K1."  
RL DNA Res. 6:83-101(1999).  
DR EMBL; AP000058; BAA79178.1;  
DR HSSP; P00782; 2SBT.  
DR InterPro; IPR000209; Peptidase\_S8.  
DR Pfam; PF00082; Peptidase\_S8; 1.  
DR PRINTS; PR00723; SUBTILISIN.  
DR PROSITE; PS00136; SUBTILASE\_ASP; 1.  
DR PROSITE; PS00137; SUBTILASE\_HIS; 1.  
DR PROSITE; PS00138; SUBTILASE\_SER; 1.  
KW Protease; Complete proteome.  
SQ SEQUENCE 440 AA; 44680 MW; E43E6E2174B6F07E CRC64;

Query Match 64.8%; Score 35; DB 17; Length 440;  
Best Local Similarity 66.7%; Pred. No. 63;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 EVVPXGMDY 10  
I::I I: I  
Db 120 EVLPWGVY 128

RESULT 18  
O52680 PRELIMINARY; PRT; 511 AA.  
ID O52680  
AC O52680; Q47289;  
DT 01-JUN-1998 (TrEMBLrel. 06, Created)  
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE Core protein (Fragment).  
GN RHSG.  
OS Escherichia coli.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Escherichia.  
OX NCBI\_TaxID=562;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ECOR-50;  
RA Wang Y.-D., Zhao S., Hill C.W.;

RT "rhs elements comprise three subfamilies.";  
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF044506; AAC62388.1; -  
FT NON\_TER 511 511  
SQ SEQUENCE 511 AA; 55877 MW; C85402569450DBFF CRC64;  
  
Query Match 64.8%; Score 35; DB 2; Length 511;  
Best Local Similarity 60.0%; Pred. No. 75;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
  
QY 1 EEVVPXGMDY 10  
Db 381 EQVNPGLDY 390  
|:|:|:|:|  
  
RESULT 19  
Q8XZL5 PRELIMINARY; PRT; 517 AA.  
AC Q8XZL5;  
DT 01-MAR-2002 (TREMBLrel. 20, Created)  
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE Probable SUBSTATE-binding periplasmic (PBP) ABC transporter protein.  
DE RSC1380 OR RS04663.  
GN Ralstonia solanacearum (Pseudomonas solanacearum).  
OS Bacteria; Proteobacteria; beta subdivision; Ralstonia group;  
OC Ralstonia.  
OC NCBI\_TaxID=305;  
OX [1]  
RN SEQUENCE FROM N.A.  
RP STRAIN-GM11000;  
RC MEDLINE=21681879; PubMed=11823852;  
RX Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S., Arlat M., Billault A., Brottier P., Camus J.C., Cattolico L., Chandlier M., Choise N., Claudel-Renard C., Cunnac S., Demange N., Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schiex T., Siguer P., Thebault P., Whalen M., Wincker P., Levy M., Weissbach J., Boucher C.A.;  
RA "Genome sequence of the plant pathogen Ralstonia solanacearum.";  
RT Nature 415:497-502(2002).  
RL EMBL; AL646064; CAD15082.1; -  
DR InterPro; IPR000914; SBP\_bac\_5.  
DR Pfam; PF00496; SBP\_bac\_5; 1.  
KW Complete proteome.  
SQ SEQUENCE 517 AA; 57349 MW; 861E836F645F2DA CRC64;  
  
Query Match 64.8%; Score 35; DB 16; Length 517;  
Best Local Similarity 66.7%; Pred. No. 76;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
  
QY 3 VVPXGMDYS 11  
Db 316 VVPQGVDA 324  
|:|:|:|:|  
  
RESULT 20  
Q95P46 PRELIMINARY; PRT; 745 AA.  
ID Q95P46;  
AC Q95P46;  
DT 01-DEC-2001 (TREMBLrel. 19, Created)  
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE Na+/K+/2Cl-cotransporter (Fragment).  
OS Carcinus maenas (Common shore crab) (Green crab).  
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;  
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;  
OC Eubranchyura; Portunoidae; Portunidae; Carcinus.  
OX NCBI\_TaxID=6759;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=GILL;  
RA Weihrauch D., Towle D.W.;

RT "Sequence and expression analysis of the Na+/K+/2Cl- cotransporter in the euryhaline crabs Callinectes sapidus and Carcinus maenas.";  
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY035548; AAK62044.1; -  
FT InterPro; IPR002293; AA/rel\_primeasel.  
DR InterPro; IPR004841; Permease.  
DR Pfam; PF00324; aa\_permeases; 1.  
FT NON\_TER 745 745  
SQ SEQUENCE 745 AA; 81249 MW; 08AE2D3E1F70C7BE CRC64;  
  
Query Match 64.8%; Score 35; DB 5; Length 745;  
Best Local Similarity 75.0%; Pred. No. 1.1e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
QY 4 VPXGMDYS 11  
Db 625 VPQGLDYS 632  
|:|:|:|:|  
  
RESULT 21  
Q9UGA3 PRELIMINARY; PRT; 1031 AA.  
ID Q9UGA3;  
AC Q9UGA3;  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE Na+/K+/2Cl-cotransporter.  
GN NKCC.  
OS Callinectes sapidus (Blue crab).  
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;  
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;  
OC Eubranchyura; Portunoidae; Portunidae; Callinectes.  
OX NCBI\_TaxID=6763;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=GILL;  
RA Towle D.W.;  
RT "Cloning and sequencing a Na+/K+/2Cl- cotransporter from gills of the euryhaline blue crab Callinectes sapidus.";  
RL Am. Zoologist 38:114A-114A(1998).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=GILL;  
RA Towle D.W., Weihrauch D.;  
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF190129; AAF05702.1; -  
DR InterPro; IPR002293; AA/rel\_primeasel.  
DR InterPro; IPR004842; KCL\_cotransport.  
DR PRINTS; PR01207; NAKCLTRNSPRT.  
DR TIGRfams; TIGR00930; 2a30; 1.  
SQ SEQUENCE 1031 AA; 113699 MW; F67773021AB05D71 CRC64;  
  
Query Match 64.8%; Score 35; DB 5; Length 1031;  
Best Local Similarity 75.0%; Pred. No. 1.7e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
QY 4 VPXGMDYS 11  
Db 737 VPQGLDYS 744  
|:|:|:|:|  
  
RESULT 22  
Q17704 PRELIMINARY; PRT; 1150 AA.  
ID Q17704  
AC Q17704; O17976;  
DT 01-JAN-1998 (TREMBLrel. 05, Created)  
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE C53A5.2 protein.  
GN C53A5.2.  
OS Caenorhabditis elegans.

Best Local Similarity 60.0%; Pred. NO. 2.4e+02;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMDY 10  
|:| | |:  
Db 382 EQVNPGLDY 391

RESULT 24  
052666 PRELIMINARY; PRT; 1420 AA.  
ID OS2666  
AC OS2666;  
DT 01-JUN-1998 (TREMBLrel. 06, Created)  
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE Core protein.  
OS Escherichia coli.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Escherichia.  
OX NCBI\_TaxID=562;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=EC45;  
RX MEDLINE=98361897; PubMed=9696756;  
RX "Wang Y.D., Zhao S., Hill C.W.;  
RT "Rhs elements comprise three subfamilies which diverged prior to  
RT acquisition by Escherichia coli.";  
RL J. Bacteriol. 180:4102-4110(1998).  
DR EMBL; AF044501; AAC32471.1; -  
DR InterPro; IPR001826; Rhs.  
DR Pfam; PF03527; Rhs; 1.  
DR PRINTS; PR00394; RhsPROTEIN.  
DR PROSITE; PS00142; ZINC\_PROTEASE; UNKNOWN\_1.  
SQ SEQUENCE 1420 AA; 159215 MW; E7851E4D48740621 CRC64;

Query Match 64.8%; Score 35; DB 2; Length 1420;  
Best Local Similarity 60.0%; Pred. NO. 2.4e+02;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMDY 10  
|:| | |:  
Db 381 EQVNPGLDY 390

RESULT 25  
027146 PRELIMINARY; PRT; 1474 AA.  
ID OS27146  
AC OS27146;  
DT 01-JAN-1998 (TREMBLrel. 05, Created)  
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE Putative membrane protein.  
GN MTH1074.  
OS Methanobacterium thermoautotrophicum.  
OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;  
OC Methanobacteriaceae; Methanothermobacter.  
OX NCBI\_TaxID=187420;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=DELTA H;  
RX MEDLINE=98037514; PubMed=9371463;  
RA Smith D.R., Doucette-Stamm L.A., DeLoughery C., Lee H.-M., Dubois J.,  
RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,  
RA Harrison D., Hoang L., Keagle P., Lumm W., Pothier B., Qiu D.,  
RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,  
RA Jiواني N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,  
RA McDougall S., Shimer G., Goyal A., Pietrowski S., Church G.M.,  
RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;  
RT "Complete genome sequence of Methanobacterium thermoautotrophicum  
RT delcA: functional analysis and comparative genomics.";  
RL J. Bacteriol. 179:7135-7155(1997).  
DR EMBL; AE000878; AAB85563.1; -.

DR InterPro; IPR001434; DUF11.  
 DR Pfam; PF01345; DUF11; 9.  
 KW Complete proteome.  
 SQ SEQUENCE 1474 AA; 153713 MW; 11D4D27BD86255F0 CRC64;

Query Match 64.8%; Score 35; DB 17; Length 1474;  
 Best Local Similarity 55.6%; Pred. No. 2.5e+02;  
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 EVVPXGMDY 10  
 :|:| |:  
 Db 1238 DVLPAGLDY 1246

Search completed: June 10, 2003, 13:46:32  
 Job time : 28.7857 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2003, 13:24:45 ; Search time 31.3571 Seconds  
(without alignments)  
46.744 Million cell updates/sec

Title: us-09-909-164-9  
Perfect score: 52  
Sequence: 1 EVVVPXGMSYS 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 75 summaries

Database :			
A.Geneseq_101002.*			
1:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*		
2:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*		
3:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*		
4:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*		
5:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*		
6:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*		
7:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*		
8:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*		
9:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*		
10:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*		
11:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*		
12:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*		
13:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*		
14:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*		
15:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*		
16:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*		
17:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*		
18:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*		
19:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*		
20:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*		
21:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*		
22:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*		
23:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*		

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Query Match	Length	ID	Description
1	50	96.2	11	23	ABB80521	Hepatitis C virus
2	50	96.2	11	23	ABB80522	Hepatitis C virus
3	50	96.2	11	23	ABB80525	Hepatitis C virus
4	50	96.2	11	23	ABB80526	Hepatitis C virus
5	50	96.2	11	23	ABB80529	Hepatitis C virus
6	50	96.2	11	23	ABB80559	Hepatitis C virus
7	50	96.2	11	23	ABB80563	Hepatitis C virus
8	50	96.2	11	23	ABB80564	Hepatitis C virus
9	50	96.2	11	23	ABB80565	Hepatitis C virus
10	50	96.2	11	23	ABB80566	Hepatitis C virus

11	50	96.2	11	23	ABB80568	Hepatitis C virus
12	46	88.5	11	23	ABB80524	Hepatitis C virus
13	46	88.5	11	23	ABB80528	Hepatitis C virus
14	46	88.5	11	23	ABB80529	Hepatitis C virus
15	46	88.5	11	23	ABB80561	Hepatitis C virus
16	46	88.5	11	23	ABB80562	Hepatitis C virus
17	45	86.5	11	23	ABB80523	Hepatitis C virus
18	45	86.5	11	23	ABB80527	Hepatitis C virus
19	45	86.5	11	23	ABB80535	Hepatitis C virus
20	45	86.5	11	23	ABB80536	Hepatitis C virus
21	45	86.5	11	23	ABB80539	Hepatitis C virus
22	45	86.5	11	23	ABB80540	Hepatitis C virus
23	45	86.5	11	23	ABB80558	Hepatitis C virus
24	45	86.5	11	23	ABB80560	Hepatitis C virus
25	44	84.6	11	23	ABB80544	Hepatitis C virus
26	44	84.6	11	23	ABB80545	Hepatitis C virus
27	44	84.6	11	23	ABB80549	Hepatitis C virus
28	44	84.6	11	23	ABB80552	Hepatitis C virus
29	44	84.6	11	23	ABB80553	Hepatitis C virus
30	42	80.8	11	23	ABB80530	Hepatitis C virus
31	41	78.8	11	23	ABB80538	Hepatitis C virus
32	41	78.8	11	23	ABB80542	Hepatitis C virus
33	41	78.8	11	23	ABB80543	Hepatitis C virus
34	40	76.9	11	23	ABB80537	Hepatitis C virus
35	40	76.9	11	23	ABB80541	Hepatitis C virus
36	40	76.9	11	23	ABB80547	Hepatitis C virus
37	40	76.9	11	23	ABB80548	Hepatitis C virus
38	40	76.9	11	23	ABB80551	Hepatitis C virus
39	40	76.9	11	23	ABB80556	Hepatitis C virus
40	40	76.9	11	23	ABB80557	Hepatitis C virus
41	40	76.9	20	20	AAU76810	Novel human diago
42	40	76.9	1022	22	ABG03621	Novel human diago
43	40	76.9	1022	22	ABG05826	Novel human diago
44	40	76.9	1022	22	ABG08173	Hepatitis C virus
45	39	75.0	11	23	ABB80546	Hepatitis C virus
46	39	75.0	11	23	ABB80550	Hepatitis C virus
47	39	75.0	11	23	ABB80554	Hepatitis C virus
48	39	75.0	11	23	ABB80555	Hepatitis C virus
49	38	73.1	11	23	ABB80533	Hepatitis C virus
50	38	73.1	11	23	ABB80534	Hepatitis C virus
51	38	73.1	3472	21	AAV90913	Cenarchaeum symbio
52	37	71.2	11	23	ABB80531	Hepatitis C virus
53	37	71.2	11	23	ABB80532	Hepatitis C virus
54	36	69.2	244	21	AA112881	Murine JNK3 bindin
55	36	69.2	484	21	AA112882	Murine JNK3 bindin
56	35	67.3	11	18	AAW99288	Peptide N424 from
57	34	65.4	842	21	AAV44359	P. chrysogenum sut
58	34	65.4	947	21	AA25105	Human novel protei
59	34	65.4	1070	22	AAU14378	Pinus radiata cell
60	33	63.5	12	21	AAV83772	HCV NS3 protease s
61	33	63.5	12	21	AAV83774	HCV NS3 protease s
62	33	63.5	13	18	AAW99276	Peptide D4 from WO
63	33	63.5	13	18	AAW99277	Peptide D3 from WO
64	33	63.5	14	18	AAW99275	Peptide C0 from WO
65	33	63.5	14	18	AAW99277	Peptide 4 used in
66	33	63.5	14	18	AAW99275	Peptide 6 used in
67	33	63.5	16	18	AAW99274	Peptide D2 from WO
68	33	63.5	16	18	AAW99274	Peptide 3 used in
69	33	63.5	16	18	AAW99273	Peptide D1 from WO
70	33	63.5	18	18	AAW99273	Peptide 2 used in
71	33	63.5	18	18	AAW99273	Cleavable substrate
72	33	63.5	18	19	AAW71281	Peptide PS from WO
73	33	63.5	20	18	AAW99272	Peptide 1 used in
74	33	63.5	20	18	AAW99272	HCV NS3 protease t
75	33	63.5	20	20	AAV14511	

ALIGNMENTS

RESULT 1  
ABB80521  
ID ABB80521 standard; peptide; 11 AA.

XX AC ABB80521;  
 XX 08-OCT-2002 (first entry)  
 XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #1.  
 DE Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
 KW virucide.  
 XX Synthetic.  
 OS Key Location/Qualifiers  
 XX Modified-site 1  
 FT /note= "N-terminal acetyl"  
 FT Modified-site 6  
 FT /note= "Norvalyl carbonyl forming keto-amide linkage with  
 FT residue 7"  
 FT Misc-difference 9  
 FT /note= "D-form residue"  
 FT Modified-site 11  
 FT /note= "C-terminal amide"  
 XX WO200208251-A2.  
 XX 31-JAN-2002.  
 PD 19-JUL-2001; 2001WO-US23169.  
 FT 21-JUL-2000; 2000US-220101P.  
 FT (CORV-) CORVAS INT INC.  
 XX Lim-wilby M, Levy OE, Brunck TK;  
 PI WPI; 2002-361643/39.  
 XX Novel peptide compound having hepatitis C virus protease inhibitory  
 PT activity useful for treating disorders associated with hepatitis C  
 PT virus protease  
 XX Claim 17; Page 64; 69pp; English.  
 XX The sequence represents a peptide compound of the invention having  
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
 CC invention are alpha-ketoamide peptide analogues. The peptides have  
 CC virucide activity, and are useful for treating and in the manufacture of  
 CC a medicament to treat disorders associated with HCV protease. A  
 CC pharmaceutical composition comprising the peptide as an active ingredient  
 CC is useful for treating disorders associated with hepatitis C virus.  
 XX Sequence 11 AA;  
 SQ Query Match 96.2%; Score 50; DB 23; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 0.0011;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 EEVVPXGMSYS 11  
 Db 1 EEVVPXGMSYS 11  
 RESULT 2  
 ABB80522  
 ID ABB80522 standard; peptide; 11 AA.  
 XX AC ABB80522;  
 XX 08-OCT-2002 (first entry)  
 XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #2.  
 DE Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
 KW virucide.  
 XX Synthetic.  
 OS Key Location/Qualifiers  
 XX Modified-site 1  
 FT /note= "N-terminal acetyl"  
 FT Modified-site 6  
 FT /note= "Norvalyl carbonyl forming keto-amide linkage with  
 FT residue 7"

OS Synthetic.  
 XX Key Location/Qualifiers  
 FT Modified-site 1  
 FT /note= "N-terminal acetyl"  
 FT Modified-site 6  
 FT /note= "Norvalyl carbonyl forming keto-amide linkage with  
 FT residue 7"  
 FT Misc-difference 9  
 FT /note= "D-form residue"  
 FT Modified-site 11  
 FT /note= "C-terminal amide"  
 XX WO200208251-A2.  
 XX 31-JAN-2002.  
 PD 19-JUL-2001; 2001WO-US23169.  
 FT 21-JUL-2000; 2000US-220101P.  
 FT (CORV-) CORVAS INT INC.  
 XX Lim-wilby M, Levy OE, Brunck TK;  
 PI WPI; 2002-361643/39.  
 XX Novel peptide compound having hepatitis C virus protease inhibitory  
 PT activity useful for treating disorders associated with hepatitis C  
 PT virus protease  
 XX Claim 17; Page 64; 69pp; English.  
 XX The sequence represents a peptide compound of the invention having  
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
 CC invention are alpha-ketoamide peptide analogues. The peptides have  
 CC virucide activity, and are useful for treating and in the manufacture of  
 CC a medicament to treat disorders associated with HCV protease. A  
 CC pharmaceutical composition comprising the peptide as an active ingredient  
 CC is useful for treating disorders associated with hepatitis C virus.  
 XX Sequence 11 AA;  
 SQ Query Match 96.2%; Score 50; DB 23; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 0.0011;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 EEVVPXGMSYS 11  
 Db 1 EEVVPXGMSYS 11  
 RESULT 3  
 ABB80525  
 ID ABB80525 standard; peptide; 11 AA.  
 XX AC ABB80525;  
 XX 08-OCT-2002 (first entry)  
 XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #5.  
 DE Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
 KW virucide.  
 XX Synthetic.  
 OS Key Location/Qualifiers  
 XX Modified-site 1  
 FT /note= "N-terminal acetyl"  
 FT Modified-site 6  
 FT /note= "Norvalyl carbonyl forming keto-amide linkage with  
 FT residue 7"





XX PA (CORV-) CORVAS INT INC.  
XX PI Lim-wilby M, Levy OE, Brunck TK;  
XX WPI; 2002-361643/39.  
XX Novel peptide compound having hepatitis C virus protease inhibitory  
PT activity useful for treating disorders associated with hepatitis C  
PT virus protease -  
XX Claim 17; Page 65; 69pp; English.  
XX PS Sequence 11 AA;  
XX CC The sequence represents a peptide compound of the invention having  
CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
CC invention are alpha-ketoamide peptide analogues. The peptides have  
CC virucide activity, and are useful for treating and in the manufacture of  
CC a medicament to treat disorders associated with HCV protease. A  
CC pharmaceutical composition comprising the peptide as an active ingredient  
CC is useful for treating disorders associated with hepatitis C virus.  
XX SQ

Query Match 96.2%; Score 50; DB 23; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.0011;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 EEVVPXGMSYS 11  
Db 1 EEVVPXGMSYS 11

RESULT 6  
ABB80563  
ID ABB80563 standard; peptide; 11 AA.  
XX AC ABB80563;  
XX DT 08-OCT-2002 (first entry)  
XX DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #43.  
XX KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
XX KW virucide.  
XX OS Synthetic.  
XX EH Key Location/Qualifiers  
FT Modified-site 1 /note= "N-terminal acetyl"  
FT Modified-site 6  
FT Modified-site 6 /note= "Valyl carbonyl forming keto-amide linkage with  
FT residue 7"  
FT Modified-site 11 /note= "C-terminal amide"  
XX WO200208251-A2.  
XX 31-JAN-2002.  
XX 19-JUL-2001; 2001WO-US23169.  
XX 21-JUL-2000; 2000US-220101P.  
XX (CORV-) CORVAS INT INC.  
XX Lim-wilby M, Levy OE, Brunck TK;  
XX WPI; 2002-361643/39.  
XX Novel peptide compound having hepatitis C virus protease inhibitory  
PT activity useful for treating disorders associated with hepatitis C  
PT virus protease -

XX PS Claim 17; Page 65; 69pp; English.  
XX CC The sequence represents a peptide compound of the invention having  
CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
CC invention are alpha-ketoamide peptide analogues. The peptides have  
CC virucide activity, and are useful for treating and in the manufacture of  
CC a medicament to treat disorders associated with HCV protease. A  
CC pharmaceutical composition comprising the peptide as an active ingredient  
CC is useful for treating disorders associated with hepatitis C virus.  
XX SQ

Query Match 96.2%; Score 50; DB 23; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.0011;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 EEVVPXGMSYS 11  
Db 1 EEVVPXGMSYS 11

RESULT 7  
ABB80564  
ID ABB80564 standard; peptide; 11 AA.  
XX AC ABB80564;  
XX DT 08-OCT-2002 (first entry)  
XX DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #44.  
XX KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
XX KW virucide.  
XX OS Synthetic.  
XX EH Key Location/Qualifiers  
FT Modified-site 1 /note= "N-terminal acetyl"  
FT Modified-site 6  
FT Modified-site 6 /note= "Leucyl carbonyl forming keto-amide linkage with  
FT residue 7"  
FT Modified-site 11 /note= "C-terminal amide"  
XX WO200208251-A2.  
XX 31-JAN-2002.  
XX 19-JUL-2001; 2001WO-US23169.  
XX 21-JUL-2000; 2000US-220101P.  
XX (CORV-) CORVAS INT INC.  
XX Lim-wilby M, Levy OE, Brunck TK;  
XX WPI; 2002-361643/39.  
XX Novel peptide compound having hepatitis C virus protease inhibitory  
PT activity useful for treating disorders associated with hepatitis C  
PT virus protease -  
XX Claim 17; Page 65; 69pp; English.  
XX CC The sequence represents a peptide compound of the invention having  
CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
CC invention are alpha-ketoamide peptide analogues. The peptides have  
CC virucide activity, and are useful for treating and in the manufacture of  
CC a medicament to treat disorders associated with HCV protease. A  
CC pharmaceutical composition comprising the peptide as an active ingredient  
CC is useful for treating disorders associated with hepatitis C virus.

```

XX SQ Sequence 11 AA;
Query Match 96.2%; Score 50; DB 23; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.0011;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEVVPXGMSYS 11
   |||||
Db 1 EEVVPXGMSYS 11

RESULT 8
ABB80565
ID ABB80565 standard; peptide; 11 AA.
XX AC ABB80565;
XX 08-OCT-2002 (first entry)
XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #45.
XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
XX virucide.
XX Synthetic.
XX Key Location/Qualifiers
XX Modified-site 1 /note= "N-terminal acetyl"
XX Modified-site 6 /note= "N-terminal acetyl"
XX Modified-site 6 /note= "N-terminal acetyl"
XX Modified-site 11 /note= "C-terminal amide"
XX WO200208251-A2.
XX 31-JAN-2002.
XX 19-JUL-2001; 2001WO-US23169.
XX 21-JUL-2000; 2000US-220101P.
XX (CORV-) CORVAS INT INC.
XX Lim-wilby M, Levy OE, Brunck TK;
XX WPI; 2002-361643/39.
XX Novel peptide compound having hepatitis C virus protease inhibitory
XX activity useful for treating disorders associated with hepatitis C
XX virus protease.
XX Claim 17; Page 65; 69pp; English.
XX The sequence represents a peptide compound of the invention having
XX hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
XX invention are alpha-ketoamide peptide analogues. The peptides have
XX virucide activity, and are useful for treating and in the manufacture of
XX a medicament to treat disorders associated with HCV protease. A
XX pharmaceutical composition comprising the peptide as an active ingredient
XX is useful for treating disorders associated with hepatitis C virus.

XX SQ Sequence 11 AA;
Query Match 96.2%; Score 50; DB 23; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.0011;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEVVPXGMSYS 11
   |||||
Db 1 EEVVPXGMSYS 11

RESULT 9
ABB80566
ID ABB80566 standard; peptide; 11 AA.
XX AC ABB80566;
XX 08-OCT-2002 (first entry)
XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #46.
XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
XX virucide.
XX Synthetic.
XX Key Location/Qualifiers
XX Modified-site 1 /note= "N-terminal acetyl"
XX Modified-site 6 /note= "2-aminoisobutyl carbonyl residue forming a
XX Modified-site 11 /note= "keto-amide linkage with residue 7"
XX WO200208251-A2.
XX 31-JAN-2002.
XX 19-JUL-2001; 2001WO-US23169.
XX 21-JUL-2000; 2000US-220101P.
XX (CORV-) CORVAS INT INC.
XX Lim-wilby M, Levy OE, Brunck TK;
XX WPI; 2002-361643/39.
XX Novel peptide compound having hepatitis C virus protease inhibitory
XX activity useful for treating disorders associated with hepatitis C
XX virus protease.
XX Claim 17; Page 65; 69pp; English.
XX The sequence represents a peptide compound of the invention having
XX hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
XX invention are alpha-ketoamide peptide analogues. The peptides have
XX virucide activity, and are useful for treating and in the manufacture of
XX a medicament to treat disorders associated with HCV protease. A
XX pharmaceutical composition comprising the peptide as an active ingredient
XX is useful for treating disorders associated with hepatitis C virus.

XX SQ Sequence 11 AA;
Query Match 96.2%; Score 50; DB 23; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.0011;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEVVPXGMSYS 11
   |||||
Db 1 EEVVPXGMSYS 11

RESULT 10
ABB80567
ID ABB80567 standard; peptide; 11 AA.
XX AC ABB80567;
XX 08-OCT-2002 (first entry)
XX

```

DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #47.  
 KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
 KW virucide.  
 XX Synthetic.  
 OS  
 XX Key Location/Qualifiers  
 FH Modified-site 1  
 FT /note= "N-terminal acetyl"  
 FT Modified-site 6  
 FT /note= "(s,s)allothreonyl carbonyl residue forming a  
 FT keto-amide linkage with residue 7"  
 FT Modified-site 11  
 FT /note= "C-terminal amide"  
 FT  
 XX WO200208251-A2.  
 PN  
 XX 31-JAN-2002.  
 PD  
 XX 19-JUL-2001; 2001WO-US23169.  
 PF  
 XX 21-JUL-2000; 2000US-220101P.  
 PR  
 XX (CORV-) CORVAS INT INC.  
 PA  
 XX Lim-wilby M, Levy OE, Brunck TK;  
 PI WPI; 2002-361643/39.  
 DR  
 XX Novel peptide compound having hepatitis C virus protease inhibitory  
 PT activity useful for treating disorders associated with hepatitis C  
 PT virus protease -  
 PT  
 XX Claim 17; Page 65; 69pp; English.  
 PS  
 XX The sequence represents a peptide compound of the invention having  
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
 CC invention are alpha-ketoamide peptide analogues. The peptides have  
 CC virucide activity, and are useful for treating and in the manufacture of  
 CC a medicament to treat disorders associated with HCV protease. A  
 CC pharmaceutical composition comprising the peptide as an active ingredient  
 CC is useful for treating disorders associated with hepatitis C virus.  
 XX  
 XX Sequence 11 AA;  
 SQ  
 Query Match 96.2%; Score 50; DB 23; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 0.0011;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 EEVVPXGMSYS 11  
 Db | | | | | | | | | |  
 1 EEVVPXGMSYS 11  
 RESULT 11  
 ABB80568  
 ID ABB80568 standard; peptide; 11 AA.  
 XX  
 AC ABB80568;  
 AC  
 DT 08-OCT-2002 (first entry)  
 DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #48.  
 XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
 KW virucide.  
 KW Synthetic.  
 OS  
 XX Key Location/Qualifiers  
 FH Modified-site 1  
 FT /note= "N-terminal acetyl"

FT Modified-site 6  
 FT /note= "Alpha-propynyl-glycyl-carbonyl residue forming  
 FT a keto-amide linkage with residue 7"  
 FT Modified-site 11  
 FT /note= "C-terminal amide"  
 XX  
 PN WO200208251-A2.  
 XX  
 XX 31-JAN-2002.  
 PD  
 XX 19-JUL-2001; 2001WO-US23169.  
 PF  
 XX 21-JUL-2000; 2000US-220101P.  
 PR  
 XX (CORV-) CORVAS INT INC.  
 PA  
 XX Lim-wilby M, Levy OE, Brunck TK;  
 PI WPI; 2002-361643/39.  
 DR  
 XX Novel peptide compound having hepatitis C virus protease inhibitory  
 PT activity useful for treating disorders associated with hepatitis C  
 PT virus protease -  
 PT  
 XX Claim 17; Page 65; 69pp; English.  
 PS  
 XX The sequence represents a peptide compound of the invention having  
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
 CC invention are alpha-ketoamide peptide analogues. The peptides have  
 CC virucide activity, and are useful for treating and in the manufacture of  
 CC a medicament to treat disorders associated with HCV protease. A  
 CC pharmaceutical composition comprising the peptide as an active ingredient  
 CC is useful for treating disorders associated with hepatitis C virus.  
 XX  
 XX Sequence 11 AA;  
 SQ  
 Query Match 96.2%; Score 50; DB 23; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 0.0011;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 EEVVPXGMSYS 11  
 Db | | | | | | | | | |  
 1 EEVVPXGMSYS 11  
 RESULT 12  
 ABB80524  
 ID ABB80524 standard; peptide; 11 AA.  
 XX  
 AC ABB80524;  
 AC  
 DT 08-OCT-2002 (first entry)  
 DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #4.  
 XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
 KW virucide.  
 KW Synthetic.  
 OS  
 XX Key Location/Qualifiers  
 FH Modified-site 1  
 FT /note= "N-terminal acetyl"  
 FT Modified-site 6  
 FT /note= "Norvalyl carbonyl forming keto-amide linkage with  
 FT residue 7"  
 FT Misc-difference 9  
 FT /note= "D-form residue"  
 FT Modified-site 11  
 FT /note= "C-terminal amide"  
 FT  
 XX WO200208251-A2.  
 PN  
 XX

us-09-909-164-9.rag

Wed Jun 11 15:45:00 2003

PI Lim-wilby M, Levy OE, Brunck TK;  
 XX WPI; 2002-361643/39.  
 XX Novel peptide compound having hepatitis C virus protease inhibitory  
 PT activity useful for treating disorders associated with hepatitis C  
 PT virus protease  
 XX  
 XX Claim 17; Page 64; 69pp; English.  
 XX  
 CC The sequence represents a peptide compound of the invention having  
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
 CC invention are alpha-ketoamide peptide analogues. The peptides have  
 CC virucide activity, and are useful for treating and in the manufacture of  
 CC a medicament to treat disorders associated with HCV protease. A  
 CC pharmaceutical composition comprising the peptide as an active ingredient  
 CC is useful for treating disorders associated with hepatitis C virus.  
 XX  
 SQ Sequence 11 AA;  
 Query Match 88.5%; Score 46; DB 23; Length 11;  
 Best Local Similarity 90.9%; Pred. No. 0.0075;  
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 EEVVPXGMSYS 11  
 |||||  
 DB 1 EEVVPXGMDYS 11

RESULT 14  
 ABB80529  
 ID ABB80529 standard; peptide; 11 AA.  
 XX  
 AC ABB80529;  
 XX  
 DT 08-OCT-2002 (first entry)  
 XX  
 DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #9.  
 XX  
 KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
 KW virucide.  
 XX  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Modified-site 1 /note= "N-terminal acetyl"  
 FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with  
 FT residue 7"  
 FT Misc-difference 8 /note= "D-form residue"  
 FT Misc-difference 9 /note= "D-form residue"  
 FT Modified-site 11 /note= "C-terminal amide"  
 FT  
 XX WO200208251-A2.  
 XX  
 XX 31-JAN-2002.  
 XX  
 XX 19-JUL-2001; 2001WO-US23169.  
 XX  
 XX 21-JUL-2000; 2000US-220101P.  
 XX  
 XX (CORV-) CORVAS INT INC.  
 XX  
 XX Lim-wilby M, Levy OE, Brunck TK;  
 XX WPI; 2002-361643/39.  
 XX Novel peptide compound having hepatitis C virus protease inhibitory  
 PT activity useful for treating disorders associated with hepatitis C

PD 31-JAN-2002.  
 PF 19-JUL-2001; 2001WO-US23169.  
 XX  
 XX 21-JUL-2000; 2000US-220101P.  
 XX  
 XX (CORV-) CORVAS INT INC.  
 XX  
 XX Lim-wilby M, Levy OE, Brunck TK;  
 XX WPI; 2002-361643/39.  
 XX Novel peptide compound having hepatitis C virus protease inhibitory  
 PT activity useful for treating disorders associated with hepatitis C  
 PT virus protease  
 XX  
 XX Claim 17; Page 64; 69pp; English.  
 XX  
 CC The sequence represents a peptide compound of the invention having  
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
 CC invention are alpha-ketoamide peptide analogues. The peptides have  
 CC virucide activity, and are useful for treating and in the manufacture of  
 CC a medicament to treat disorders associated with HCV protease. A  
 CC pharmaceutical composition comprising the peptide as an active ingredient  
 CC is useful for treating disorders associated with hepatitis C virus.  
 XX  
 SQ Sequence 11 AA;  
 Query Match 88.5%; Score 46; DB 23; Length 11;  
 Best Local Similarity 90.9%; Pred. No. 0.0075;  
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 EEVVPXGMSYS 11  
 |||||  
 DB 1 EEVVPXGMDYS 11

RESULT 13  
 ABB80528  
 ID ABB80528 standard; peptide; 11 AA.  
 XX  
 AC ABB80528;  
 XX  
 DT 08-OCT-2002 (first entry)  
 XX  
 DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #8.  
 XX  
 KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
 KW virucide.  
 XX  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Modified-site 1 /note= "N-terminal acetyl"  
 FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with  
 FT residue 7"  
 FT Misc-difference 8 /note= "D-form residue"  
 FT Modified-site 11 /note= "C-terminal amide"  
 FT  
 XX WO200208251-A2.  
 XX  
 XX 31-JAN-2002.  
 XX  
 XX 19-JUL-2001; 2001WO-US23169.  
 XX  
 XX 21-JUL-2000; 2000US-220101P.  
 XX  
 XX (CORV-) CORVAS INT INC.  
 XX

PT virus protease -  
XX Claim 17; Page 64; 69pp; English.  
XX  
CC The sequence represents a peptide compound of the invention having  
CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
CC invention are alpha-ketoamide peptide analogues. The peptides have  
CC virucide activity, and are useful for treating and in the manufacture of  
CC a medicament to treat disorders associated with HCV protease. A  
CC pharmaceutical composition comprising the peptide as an active ingredient  
CC is useful for treating disorders associated with hepatitis C virus.  
XX  
SQ Sequence 11 AA;  
  
Query Match 88.5%; Score 46; DB 23; Length 11;  
Best Local Similarity 90.9%; Pred. No. 0.0075;  
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 EEVVPXGMSYS 11  
Db 1 EEVVPXGMDYS 11  
IIIIIIII II  
  
RESULT 15  
ID ABB80561 standard; peptide; 11 AA.  
XX  
AC ABB80561;  
XX  
DT 08-OCT-2002 (first entry)  
XX  
DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #41.  
XX  
KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
KW virucide.  
XX  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT Modified-site 1 /note= "N-terminal acetyl"  
FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with  
FT Misc-difference 8 /note= "D-form residue"  
FT Modified-site 8 /note= "Oxymethionine"  
FT Misc-difference 8 /note= "D-form residue"  
FT Modified-site 11 /note= "C-terminal amide"  
FT Modified-site 11 /note= "C-terminal amide"  
XX  
PN WO200208251-A2.  
XX  
PD 31-JAN-2002.  
XX  
PF 19-JUL-2001; 2001WO-US23169.  
XX  
PR 21-JUL-2000; 2000US-220101P.  
XX  
PA (CORV-) CORVAS INT INC.  
XX  
PI Lim-wilby M, Levy OE, Brunck TK;  
XX  
DR WPI; 2002-361643/39.  
XX  
PT Novel peptide compound having hepatitis C virus protease inhibitory  
PT activity useful for treating disorders associated with hepatitis C  
PT virus protease  
XX  
PS Claim 17; Page 65; 69pp; English.  
XX  
CC The sequence represents a peptide compound of the invention having  
CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the

CC invention are alpha-ketoamide peptide analogues. The peptides have  
CC virucide activity, and are useful for treating and in the manufacture of  
CC a medicament to treat disorders associated with HCV protease. A  
CC pharmaceutical composition comprising the peptide as an active ingredient  
CC is useful for treating disorders associated with hepatitis C virus.  
XX  
SQ Sequence 11 AA;  
  
Query Match 88.5%; Score 46; DB 23; Length 11;  
Best Local Similarity 90.9%; Pred. No. 0.0075;  
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 EEVVPXGMSYS 11  
Db 1 EEVVPXGMDYS 11  
IIIIIIII II  
  
RESULT 16  
ID ABB80562 standard; peptide; 11 AA.  
XX  
AC ABB80562;  
XX  
DT 08-OCT-2002 (first entry)  
XX  
DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #42.  
XX  
KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
KW virucide.  
XX  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT Modified-site 1 /note= "N-terminal acetyl"  
FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with  
FT Misc-difference 8 /note= "D-form residue"  
FT Modified-site 8 /note= "Oxymethionine"  
FT Misc-difference 9 /note= "D-form residue"  
FT Modified-site 11 /note= "C-terminal amide"  
FT Modified-site 11 /note= "C-terminal amide"  
XX  
PN WO200208251-A2.  
XX  
PD 31-JAN-2002.  
XX  
PF 19-JUL-2001; 2001WO-US23169.  
XX  
PR 21-JUL-2000; 2000US-220101P.  
XX  
PA (CORV-) CORVAS INT INC.  
XX  
PI Lim-wilby M, Levy OE, Brunck TK;  
XX  
DR WPI; 2002-361643/39.  
XX  
PT Novel peptide compound having hepatitis C virus protease inhibitory  
PT activity useful for treating disorders associated with hepatitis C  
PT virus protease  
XX  
PS Claim 17; Page 65; 69pp; English.  
XX  
CC The sequence represents a peptide compound of the invention having  
CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
CC invention are alpha-ketoamide peptide analogues. The peptides have  
CC virucide activity, and are useful for treating and in the manufacture of  
CC a medicament to treat disorders associated with HCV protease. A  
CC pharmaceutical composition comprising the peptide as an active ingredient

CC is useful for treating disorders associated with hepatitis C virus.

XX Sequence 11 AA;

Query Match 88.5%; Score 46; DB 23; Length 11;  
Best Local Similarity 90.9%; Pred. No. 0.0075;  
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXGMSYS 11  
| | | | | | | | | |  
DB 1 EEVVPXGMDYS 11

# RESULT 17

ABB80523  
ID ABB80523 standard; peptide; 11 AA.

XX ABB80523;

XX 08-OCT-2002 (first entry)

XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #3.

XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
virucide.

XX Synthetic.

XX Key Location/Qualifiers

FT Modified-site 1 /note= "N-terminal acetyl"

FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with  
residue 7"

FT Misc-difference 9 /note= "D-form residue"

FT Modified-site 11 /note= "C-terminal amide"

FT WO200208251-A2.

XX 31-JAN-2002.

XX 19-JUL-2001; 2001WO-US23169.

XX 21-JUL-2000; 2000US-220101P.

XX (CORV-) CORVAS INT INC.

XX Lim-wilby M, Levy OE, Brunck TK;

XX WPI; 2002-361643/39.

XX Novel peptide compound having hepatitis C virus protease inhibitory  
activity useful for treating disorders associated with hepatitis C  
virus protease -

XX Claim 17; Page 64; 69pp; English.

XX The sequence represents a peptide compound of the invention having  
hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
invention are alpha-ketoamide peptide analogues. The peptides have  
virucide activity, and are useful for treating and in the manufacture of  
a medicament to treat disorders associated with HCV protease. A  
pharmaceutical composition comprising the peptide as an active ingredient  
is useful for treating disorders associated with hepatitis C virus.

XX Sequence 11 AA;

Query Match 86.5%; Score 45; DB 23; Length 11;

Best Local Similarity 90.9%; Pred. No. 0.012;  
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXGMSYS 11  
| | | | | | | | | |  
DB 1 EEVVPXGMHYS 11

# RESULT 18

ABB80527  
ID ABB80527 standard; peptide; 11 AA.

XX ABB80527;

XX 08-OCT-2002 (first entry)

XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #7.

XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
virucide.

XX Synthetic.

XX Key Location/Qualifiers

FT Modified-site 1 /note= "N-terminal acetyl"

FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with  
residue 7"

FT Misc-difference 8 /note= "D-form residue"

FT Modified-site 11 /note= "C-terminal amide"

FT WO200208251-A2.

XX 31-JAN-2002.

XX 19-JUL-2001; 2001WO-US23169.

XX 21-JUL-2000; 2000US-220101P.

XX (CORV-) CORVAS INT INC.

XX Lim-wilby M, Levy OE, Brunck TK;

XX WPI; 2002-361643/39.

XX Novel peptide compound having hepatitis C virus protease inhibitory  
activity useful for treating disorders associated with hepatitis C  
virus protease -

XX Claim 17; Page 64; 69pp; English.

XX The sequence represents a peptide compound of the invention having  
hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
invention are alpha-ketoamide peptide analogues. The peptides have  
virucide activity, and are useful for treating and in the manufacture of  
a medicament to treat disorders associated with HCV protease. A  
pharmaceutical composition comprising the peptide as an active ingredient  
is useful for treating disorders associated with hepatitis C virus.

XX Sequence 11 AA;

Query Match 86.5%; Score 45; DB 23; Length 11;

Best Local Similarity 90.9%; Pred. No. 0.012;  
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXGMSYS 11  
| | | | | | | | | |  
DB 1 EEVVPXGMHYS 11

# RESULT 19

ABB80535  
ID ABB80535 standard; peptide; 11 AA.

XX AC ABB80535;  
XX DT 08-OCT-2002 (first entry)  
XX DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #15.  
XX KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
XX KW virucide.  
XX OS Synthetic.  
XX FH Key Location/Qualifiers  
XX FT Modified-site 1 /note= "N-terminal acetyl"  
XX FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with  
XX FT residue 7"  
XX FT Misc-difference 9 /note= "D-form residue"  
XX FT Modified-site 11 /note= "C-terminal amide"  
XX PN WO200208251-A2.  
XX PD 31-JAN-2002.  
XX PF 19-JUL-2001; 2001WO-US23169.  
XX PR 21-JUL-2000; 2000US-220101P.  
XX PA (CORV-) CORVAS INT INC.  
XX PI Lim-wilby M, Levy OE, Brunck TK;  
XX DR WPI; 2002-361643/39.  
XX CC Novel peptide compound having hepatitis C virus protease inhibitory  
XX FT activity useful for treating disorders associated with hepatitis C  
XX FT virus protease  
XX PS Claim 17; Page 64; 69pp; English.  
XX CC The sequence represents a peptide compound of the invention having  
XX CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
XX CC invention are alpha-ketoamide peptide analogues. The peptides have  
XX CC virucide activity, and are useful for treating and in the manufacture of  
XX CC a medicament to treat disorders associated with HCV protease. A  
XX CC pharmaceutical composition comprising the peptide as an active ingredient  
XX CC is useful for treating disorders associated with hepatitis C virus.  
XX SQ Sequence 11 AA;  
XX Query Match 86.5%; Score 45; DB 23; Length 11;  
XX Best Local Similarity 90.9%; Pred. No. 0.012;  
XX Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 EEVVPXGMSYS 11  
DB 1 EEVVPXGMSYS 11  
RESULT 20  
ABB80536  
ID ABB80536 standard; peptide; 11 AA.  
XX AC ABB80536;  
XX DT 08-OCT-2002 (first entry)  
XX DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #16.  
XX DE Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
XX KW virucide.  
XX OS Synthetic.  
XX FH Key Location/Qualifiers  
XX FT Modified-site 1 /note= "N-terminal acetyl"  
XX FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with  
XX FT residue 7"  
XX PN WO200208251-A2.  
XX PD 31-JAN-2002.  
XX PF 19-JUL-2001; 2001WO-US23169.  
XX PR 21-JUL-2000; 2000US-220101P.  
XX PA (CORV-) CORVAS INT INC.  
XX PI Lim-wilby M, Levy OE, Brunck TK;  
XX DR WPI; 2002-361643/39.  
XX CC Novel peptide compound having hepatitis C virus protease inhibitory  
XX FT activity useful for treating disorders associated with hepatitis C  
XX FT virus protease  
XX PS Claim 17; Page 64; 69pp; English.  
XX CC The sequence represents a peptide compound of the invention having  
XX CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
XX CC invention are alpha-ketoamide peptide analogues. The peptides have  
XX CC virucide activity, and are useful for treating and in the manufacture of  
XX CC a medicament to treat disorders associated with HCV protease. A  
XX CC pharmaceutical composition comprising the peptide as an active ingredient  
XX CC is useful for treating disorders associated with hepatitis C virus.  
XX SQ Sequence 11 AA;  
XX Query Match 86.5%; Score 45; DB 23; Length 11;  
XX Best Local Similarity 90.9%; Pred. No. 0.012;  
XX Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 EEVVPXGMSYS 11  
DB 1 EEVVPXGMSYS 11  
RESULT 21  
ABB80539  
ID ABB80539 standard; peptide; 11 AA.  
XX AC ABB80539;  
XX DT 08-OCT-2002 (first entry)  
XX DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #19.  
XX DE Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
XX KW virucide.  
XX OS Synthetic.  
XX FH Key Location/Qualifiers  
XX FT Modified-site 1 /note= "N-terminal acetyl"  
XX FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with  
XX FT residue 7"  
XX PN WO200208251-A2.  
XX PD 31-JAN-2002.  
XX PF 19-JUL-2001; 2001WO-US23169.  
XX PR 21-JUL-2000; 2000US-220101P.  
XX PA (CORV-) CORVAS INT INC.  
XX PI Lim-wilby M, Levy OE, Brunck TK;  
XX DR WPI; 2002-361643/39.  
XX CC Novel peptide compound having hepatitis C virus protease inhibitory  
XX FT activity useful for treating disorders associated with hepatitis C  
XX FT virus protease  
XX PS Claim 17; Page 64; 69pp; English.  
XX CC The sequence represents a peptide compound of the invention having  
XX CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
XX CC invention are alpha-ketoamide peptide analogues. The peptides have  
XX CC virucide activity, and are useful for treating and in the manufacture of  
XX CC a medicament to treat disorders associated with HCV protease. A  
XX CC pharmaceutical composition comprising the peptide as an active ingredient  
XX CC is useful for treating disorders associated with hepatitis C virus.  
XX SQ Sequence 11 AA;  
XX Query Match 86.5%; Score 45; DB 23; Length 11;  
XX Best Local Similarity 90.9%; Pred. No. 0.012;  
XX Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 EEVVPXGMSYS 11  
DB 1 EEVVPXGMSYS 11  
RESULT 21  
ABB80539  
ID ABB80539 standard; peptide; 11 AA.  
XX AC ABB80539;  
XX DT 08-OCT-2002 (first entry)  
XX DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #19.  
XX DE Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
XX KW virucide.  
XX OS Synthetic.  
XX FH Key Location/Qualifiers  
XX FT Modified-site 1 /note= "N-terminal acetyl"  
XX FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with  
XX FT residue 7"



FT Misc-difference 8 /note= "D-form residue"  
FT Modified-site 11  
FT FT /note= "C-terminal amide"  
XX WO200208251-A2.  
XX PD  
XX 31-JAN-2002.  
XX  
XX 19-JUL-2001; 2001WO-US23169.  
XX PR  
XX 21-JUL-2000; 2000US-220101P.  
XX PA (CORV-) CORVAS INT INC.  
XX PI Lim-wilby M, Levy OE, Brunck TK;  
XX WPI; 2002-361643/39.  
XX Novel peptide compound having hepatitis C virus protease inhibitory  
PT activity useful for treating disorders associated with hepatitis C  
PT virus protease  
XX  
XX Claim 17; Page 65; 69pp; English.  
XX The sequence represents a peptide compound of the invention having  
CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
CC invention are alpha-ketoamide peptide analogues. The peptides have  
CC virucide activity, and are useful for treating and in the manufacture of  
CC a medicament to treat disorders associated with HCV protease. A  
CC pharmaceutical composition comprising the peptide as an active ingredient  
CC is useful for treating disorders associated with hepatitis C virus.  
XX  
XX Sequence 11 AA;  
Query Match 86.5%; Score 45; DB 23; Length 11;  
Best Local Similarity 90.9%; Pred. No. 0.012;  
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 EEVVPXGMSYS 11  
DB 1 EEVVPXGQSYS 11  
RESULT 22  
ABB80540  
ID ABB80540 standard; peptide; 11 AA.  
XX  
XX ABB80540;  
XX 08-OCT-2002 (first entry)  
XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #20.  
XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
XX virucide.  
XX Synthetic.  
XX Key Location/Qualifiers  
FH Modified-site 1 /note= "N-terminal acetyl"  
FT Modified-site 6  
FT Modified-site 8 /note= "Norvalyl carbonyl forming keto-amide linkage with  
FT residue 7"  
FT Misc-difference 8 /note= "D-form residue"  
FT Misc-difference 9 /note= "D-form residue"  
FT Modified-site 11 /note= "C-terminal amide"  
XX WO200208251-A2.  
XX PD  
XX 31-JAN-2002.  
XX 19-JUL-2001; 2001WO-US23169.  
XX PF  
XX 21-JUL-2000; 2000US-220101P.  
XX PR (CORV-) CORVAS INT INC.

XX 31-JAN-2002.  
PD 19-JUL-2001; 2001WO-US23169.  
XX 21-JUL-2000; 2000US-220101P.  
XX (CORV-) CORVAS INT INC.  
XX Lim-wilby M, Levy OE, Brunck TK;  
XX WPI; 2002-361643/39.  
XX Novel peptide compound having hepatitis C virus protease inhibitory  
PT activity useful for treating disorders associated with hepatitis C  
PT virus protease  
XX  
XX Claim 17; Page 65; 69pp; English.  
XX The sequence represents a peptide compound of the invention having  
CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
CC invention are alpha-ketoamide peptide analogues. The peptides have  
CC virucide activity, and are useful for treating and in the manufacture of  
CC a medicament to treat disorders associated with HCV protease. A  
CC pharmaceutical composition comprising the peptide as an active ingredient  
CC is useful for treating disorders associated with hepatitis C virus.  
XX  
XX Sequence 11 AA;  
Query Match 86.5%; Score 45; DB 23; Length 11;  
Best Local Similarity 90.9%; Pred. No. 0.012;  
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 EEVVPXGMSYS 11  
DB 1 EEVVPXGQSYS 11  
RESULT 23  
ABB80558  
ID ABB80558 standard; peptide; 11 AA.  
XX  
XX ABB80558;  
XX 08-OCT-2002 (first entry)  
XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #38.  
XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
XX virucide.  
XX Synthetic.  
XX Key Location/Qualifiers  
FH Modified-site 1 /note= "N-terminal acetyl"  
FT Modified-site 6  
FT Modified-site 8 /note= "Norvalyl carbonyl forming keto-amide linkage with  
FT residue 7"  
FT Modified-site 11 /note= "Oxymethionine"  
FT Modified-site 11 /note= "C-terminal amide"  
XX WO200208251-A2.  
XX PD  
XX 31-JAN-2002.  
XX 19-JUL-2001; 2001WO-US23169.  
XX PF  
XX 21-JUL-2000; 2000US-220101P.  
XX PR (CORV-) CORVAS INT INC.  
XX PA

XX Lim-wilby M, Levy OE, Brunck TK;  
XX WPI; 2002-361643/39.  
XX  
PT Novel peptide compound having hepatitis C virus protease inhibitory  
PT activity useful for treating disorders associated with hepatitis C  
PT virus protease  
XX  
PS Claim 17; Page 65; 69pp; English.  
XX  
CC The sequence represents a peptide compound of the invention having  
CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
CC invention are alpha-ketoamide peptide analogues. The peptides have  
CC virucide activity, and are useful for treating and in the manufacture of  
CC a medicament to treat disorders associated with HCV protease. A  
CC pharmaceutical composition comprising the peptide as an active ingredient  
CC is useful for treating disorders associated with hepatitis C virus.  
XX  
SQ Sequence 11 AA;  
Query Match 86.5%; Score 45; DB 23; Length 11;  
Best Local Similarity 90.9%; Pred. No. 0.012;  
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 EEVVPXGMSYS 11  
DB 1 EEVVPXGMHYS 11  
RESULT 24  
ABB80560  
ID ABB80560 standard; peptide; 11 AA.  
XX  
AC ABB80560;  
XX  
DT 08-OCT-2002 (first entry)  
XX  
DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #40.  
XX  
KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
KW virucide.  
XX  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT Modified-site 1 /note= "N-terminal acetyl"  
FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with  
FT Modified-site residue 7"  
FT Misc-difference 8 /note= "D-form residue"  
FT Modified-site 8 /note= "Oxymethionine"  
FT Misc-difference 9 /note= "D-form residue"  
FT Modified-site 11 /note= "C-terminal amide"  
XX  
PN WO200208251-A2.  
XX  
PD 31-JAN-2002.  
XX  
PF 19-JUL-2001; 2001WO-US23169.  
XX  
PR 21-JUL-2000; 2000US-220101P.  
XX  
PA (CORV-) CORVAS INT INC.  
XX  
PI Lim-wilby M, Levy OE, Brunck TK;  
XX  
XX WPI; 2002-361643/39.

XX Novel peptide compound having hepatitis C virus protease inhibitory  
PT activity useful for treating disorders associated with hepatitis C  
PT virus protease  
XX  
PS Claim 17; Page 65; 69pp; English.  
XX  
CC The sequence represents a peptide compound of the invention having  
CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
CC invention are alpha-ketoamide peptide analogues. The peptides have  
CC virucide activity, and are useful for treating and in the manufacture of  
CC a medicament to treat disorders associated with HCV protease. A  
CC pharmaceutical composition comprising the peptide as an active ingredient  
CC is useful for treating disorders associated with hepatitis C virus.  
XX  
SQ Sequence 11 AA;  
Query Match 86.5%; Score 45; DB 23; Length 11;  
Best Local Similarity 90.9%; Pred. No. 0.012;  
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 EEVVPXGMSYS 11  
DB 1 EEVVPXGMHYS 11  
RESULT 25  
ABB80544  
ID ABB80544 standard; peptide; 11 AA.  
XX  
AC ABB80544;  
XX  
DT 08-OCT-2002 (first entry)  
XX  
DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #24.  
XX  
KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
KW virucide.  
XX  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT Modified-site 1 /note= "N-terminal acetyl"  
FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with  
FT Modified-site residue 7"  
FT Modified-site 11 /note= "C-terminal amide"  
XX  
PN WO200208251-A2.  
XX  
PD 31-JAN-2002.  
XX  
PF 19-JUL-2001; 2001WO-US23169.  
XX  
PR 21-JUL-2000; 2000US-220101P.  
XX  
PA (CORV-) CORVAS INT INC.  
XX  
PI Lim-wilby M, Levy OE, Brunck TK;  
XX  
XX WPI; 2002-361643/39.  
XX  
PT Novel peptide compound having hepatitis C virus protease inhibitory  
PT activity useful for treating disorders associated with hepatitis C  
PT virus protease  
XX  
PS Claim 17; Page 65; 69pp; English.  
XX  
CC The sequence represents a peptide compound of the invention having  
CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
CC invention are alpha-ketoamide peptide analogues. The peptides have  
XX

Wed Jun 11 15:45:00 2003

CC virucide activity, and are useful for treating and in the manufacture of  
 CC a medicament to treat disorders associated with HCV protease. A  
 CC pharmaceutical composition comprising the peptide as an active ingredient  
 CC is useful for treating disorders associated with hepatitis C virus.

xx

SQ Sequence 11 AA;

Query Match 84.6%; Score 44; DB 23; Length 11;

Best Local Similarity 90.9%; Pred. No. 0.019;

Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXGMSYS 11

|||||

Db 1 EEVVPXGTSYS 11

Search completed: June 10, 2003, 13:39:08

Job time : 32.3571 secs



291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466	467	468	469	470	471	472	473	474	475	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520	521	522	523	524	525	526	527	528	529	530	531	532	533	534	535	536	537	538	539	540	541	542	543	544	545	546	547	548	549	550	551	552	553	554	555	556	557	558	559	560	561	562	563	564	565	566	567	568	569	570	571	572	573	574	575	576	577	578	579	580	581	582	583	584	585	586	587	588	589	590	591	592	593	594	595	596	597	598	599	600	601	602	603	604	605	606	607	608	609	610	611	612	613	614	615	616	617	618	619	620	621	622	623	624	625	626	627	628	629	630	631	632	633	634	635	636	637	638	639	640	641	642	643	644	645	646	647	648	649	650	651	652	653	654	655	656	657	658	659	660	661	662	663	664	665	666	667	668	669	670	671	672	673	674	675	676	677	678	679	680	681	682	683	684	685	686	687	688	689	690	691	692	693	694	695	696	697	698	699	700	701	702	703	704	705	706	707	708	709	710	711	712	713	714	715	716	717	718	719	720	721	722	723	724	725	726	727	728	729	730	731	732	733	734	735	736	737	738	739	740	741	742	743	744	745	746	747	748	749	750	751	752	753	754	755	756	757	758	759	760	761	762	763	764	765	766	767	768	769	770	771	772	773	774	775	776	777	778	779	780	781	782	783	784	785	786	787	788	789	790	791	792	793	794	795	796	797	798	799	800	801
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

US-09-228-986-73

5 Seconds  
ts)  
all updates/sec

GenCore	GenCore
right (c) 1993	right (c) 1993
search, using sw	search, using sw
2003, 13:11	2003, 13:11
909-164-9	909-164-9
PXGMSYS 11	PXGMSYS 11
62	62
10.0, Gapext (	10.0, Gapext (
seqs, 29422922	seqs, 29422922
satisfying chos	satisfying chos
0	0
2000000000	2000000000
um Match 0%	um Match 0%
um Match 100%	um Match 100%
first 75 sum	first 75 sum
d.Parents.AA:*	d.Parents.AA:*
gn2_6/ptodata/	gn2_6/ptodata/
gn2_6/ptodata/	gn2_6/ptodata/
gn2_6/ptodata/	gn2_6/ptodata/
gn2_6/ptodata/	gn2_6/ptodata/
gn2_6/ptodata/	gn2_6/ptodata/
gn2_6/ptodata/	gn2_6/ptodata/
number of res	number of res
than or equal	than or equal
analysis of	analysis of
Length DB ID	Length DB ID
947 4 US-	947 4 US-

\_\_\_\_\_

51  
1  
1  
30  
1-945-236  
77-952-66  
11-650-66  
88-888-66  
00-988A-23  
00-988A-23  
00-694-4  
00-744-4  
93-977-7  
97-711B-4  
94-517-21  
66-361A-21  
83-772-21  
3-05000-21  
64-517-22  
66-361A-22  
83-772-22  
3-05000-22  
64-517-6  
63-772-6  
66-361A-4  
3-05000-4

45	4	US
65	6	US
517	6	US
410	6	US
1394	6	US
10	4	US
10	4	US
10	4	US
102	2	US
152	2	US
152	3	US
152	3	US
152	3	US
173	1	US
189	2	US
189	2	US
189	3	US
189	3	US
189	5	PC
236	2	US
236	2	US
236	3	US
236	5	PC
280	2	US
280	3	US
289	2	US
289	5	PC

33	4	33
33	5	33
33	6	33
33	7	33
32	8	32
32	9	32
32	10	32
32	11	32
32	12	32
32	13	32
32	14	32
32	15	32
32	16	32
32	17	32
32	18	32
32	19	32
32	20	32
32	21	32
32	22	32
32	23	32
32	24	32
32	25	32
32	26	32
32	27	32

\_\_\_\_\_

Query Match 65.4%; Score 34; DB 4; Length 947;  
Best Local Similarity 66.7%; Pred. No. 1e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVEXGMSYS 11  
Db 686 VMPGIGSYS 694

## RESULT 2

US-08-637-759B-236  
; Sequence 236, Application US/08637759B  
; Patent No. 5876931  
; GENERAL INFORMATION:  
; APPLICANT: David William Holden  
; TITLE OF INVENTION: Identification of Genes  
; NUMBER OF SEQUENCES: 501  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Patrea L. Pabst  
; STREET: 2800 One Atlantic Center  
; CITY: Atlanta  
; STATE: Georgia  
; COUNTRY: USA  
; ZIP: 30309-3450  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/637,759B  
; FILING DATE: 03-MAY-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/GB95/02875  
; FILING DATE: 11-DEC-1995  
; CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:  
; NAME: Pabst, Patrea L.  
; REGISTRATION NUMBER: 31,284  
; REFERENCE/DOCKET NUMBER: RPMS 101  
; TELEPHONE: (404) 873-8794  
; TELEFAX: (404) 873-8795  
; INFORMATION FOR SEQ ID NO: 236:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 45 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
US-08-637-759B-236

Query Match 63.5%; Score 33; DB 2; Length 45;  
Best Local Similarity 60.0%; Pred. No. 5.1;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10  
Db 1 EEISPLGWSY 10

## RESULT 3

US-08-871-355A-236  
; Sequence 236, Application US/08871355A  
; Patent No. 6015669  
; GENERAL INFORMATION:  
; APPLICANT: David William Holden  
; TITLE OF INVENTION: Identification of Genes  
; NUMBER OF SEQUENCES: 501  
; CORRESPONDENCE ADDRESS:

ADDRESSEE: Patrea L. Pabst  
; STREET: 2800 One Atlantic Center  
; CITY: Atlanta  
; STATE: Georgia  
; COUNTRY: USA  
; ZIP: 30309-3450  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/871,355A  
; FILING DATE: 09-JUN-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/GB95/02875  
; FILING DATE: 11-DEC-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Pabst, Patrea L.  
; REGISTRATION NUMBER: 31,284  
; REFERENCE/DOCKET NUMBER: RPMS 101 CON  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (404) 873-8794  
; TELEFAX: (404) 873-8795  
; INFORMATION FOR SEQ ID NO: 236:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 45 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
US-08-871-355A-236

Query Match 63.5%; Score 33; DB 3; Length 45;  
Best Local Similarity 60.0%; Pred. No. 5.1;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10  
Db 1 EEISPLGWSY 10

## RESULT 4

US-09-201-945-236  
; Sequence 236, Application US/09201945  
; Patent No. 6342215  
; GENERAL INFORMATION:  
; APPLICANT: David William Holden  
; TITLE OF INVENTION: Identification of Genes  
; NUMBER OF SEQUENCES: 501  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Patrea L. Pabst  
; STREET: 2800 One Atlantic Center  
; CITY: Atlanta  
; STATE: Georgia  
; COUNTRY: USA  
; ZIP: 30309-3450  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/201,945  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/637,759

Query Match 63.5%; Score 33; DB 2; Length 45;  
Best Local Similarity 60.0%; Pred. No. 5.1;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10  
Db 1 EEISPLGWSY 10

## RESULT 3

US-08-871-355A-236  
; Sequence 236, Application US/08871355A  
; Patent No. 6015669  
; GENERAL INFORMATION:  
; APPLICANT: David William Holden  
; TITLE OF INVENTION: Identification of Genes  
; NUMBER OF SEQUENCES: 501  
; CORRESPONDENCE ADDRESS:

;  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Pabst, Patrea L.  
; REGISTRATION NUMBER: 31,284  
; REFERENCE/DOCKET NUMBER: RPMs 101  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (404) 873-8794  
; TELEFAX: (404) 873-8795  
; INFORMATION FOR SEQ ID NO: 236:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 45 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
; US-09-201-945-236

Query Match 63.5%; Score 33; DB 4; Length 45;  
Best Local Similarity 60.0%; Pred. No. 5.1;  
Matches 6; Conservative 1; Mismatches 3; Indels 3; Gaps 0;

Qy 1 EEVVPXGMSY 10  
||: ||| |  
Db 1 EEISPLGWSY 10

RESULT 5  
517197-51  
; Patent No. 517197  
; APPLICANT: KANZAKI, TETSUTO; OLOFSSON, ANDERS; MOREN, ANITA;  
; WERNSTEDT, CHRISTER; HELLMAN, ULF; MIYAZONO, KOHEI; CLAESSEN-WELSH,  
; LENA; HELDIN, CARL-HENRIK  
; TITLE OF INVENTION: ISOLATED NUCLEOTIDE SEQUENCE EXPRESSING  
; HUMAN TRANSFORMING GROWTH FACTOR-BETAL-BINDING PROTEIN  
; NUMBER OF SEQUENCES: 53  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/487,343  
; FILING DATE: 27-FEB-1990  
; SEQ ID NO: 51:  
; LENGTH: 65  
517197-51

Query Match 63.5%; Score 33; DB 6; Length 65;  
Best Local Similarity 45.5%; Pred. No. 7.8;  
Matches 5; Conservative 3; Mismatches 3; Indels 3; Gaps 0;

Qy 1 EEVVPXGMSYS 11  
||: ||| |  
Db 52 KEICPGMGYT 62

RESULT 6  
517197-1  
; Patent No. 517197  
; APPLICANT: KANZAKI, TETSUTO; OLOFSSON, ANDERS; MOREN, ANITA;  
; WERNSTEDT, CHRISTER; HELLMAN, ULF; MIYAZONO, KOHEI; CLAESSEN-WELSH,  
; LENA; HELDIN, CARL-HENRIK  
; TITLE OF INVENTION: ISOLATED NUCLEOTIDE SEQUENCE EXPRESSING  
; HUMAN TRANSFORMING GROWTH FACTOR-BETAL-BINDING PROTEIN  
; NUMBER OF SEQUENCES: 53  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/487,343  
; FILING DATE: 27-FEB-1990  
; SEQ ID NO: 1:  
; LENGTH: 410  
517197-1

Query Match 63.5%; Score 33; DB 6; Length 410;  
Best Local Similarity 45.5%; Pred. No. 63;  
Matches 5; Conservative 3; Mismatches 3; Indels 3; Gaps 0;

Qy 1 EEVVPXGMSYS 11  
||: ||| |  
Db 399 KEICPGMGYT 409

RESULT 7  
517197-30  
; Patent No. 517197  
; APPLICANT: KANZAKI, TETSUTO; OLOFSSON, ANDERS; MOREN, ANITA;  
; WERNSTEDT, CHRISTER; HELLMAN, ULF; MIYAZONO, KOHEI; CLAESSEN-WELSH,  
; LENA; HELDIN, CARL-HENRIK  
; TITLE OF INVENTION: ISOLATED NUCLEOTIDE SEQUENCE EXPRESSING  
; HUMAN TRANSFORMING GROWTH FACTOR-BETAL-BINDING PROTEIN  
; NUMBER OF SEQUENCES: 53  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/487,343  
; FILING DATE: 27-FEB-1990  
; SEQ ID NO: 30:  
; LENGTH: 1394  
517197-30

Query Match 63.5%; Score 33; DB 6; Length 1394;  
Best Local Similarity 45.5%; Pred. No. 2.6e+02;  
Matches 5; Conservative 3; Mismatches 3; Indels 3; Gaps 0;

Qy 1 EEVVPXGMSYS 11  
||: ||| |  
Db 399 KEICPGMGYT 409

RESULT 8  
US-09-357-952-66  
; Sequence 66, Application US/09357952  
; Patent No. 6248904  
; GENERAL INFORMATION:  
; APPLICANT: Zhang, Han-Zhong  
; APPLICANT: Cai, Sui Xiong  
; APPLICANT: Drewe, John A.  
; APPLICANT: Yang, Wu  
; TITLE OF INVENTION: No. 6248904el Fluorescence Dyes and Their Applications for  
; TITLE OF INVENTION: Fluorescence Screening Assays for Caspases, Peptidases, Pro  
; FILE REFERENCE: 1735.0030001  
; CURRENT APPLICATION NUMBER: US/09/357,952  
; CURRENT FILING DATE: 1999-07-21  
; EARLIER APPLICATION NUMBER: US 60/093,642  
; EARLIER FILING DATE: 21-JUL-1998  
; NUMBER OF SEQ ID NOS: 139  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 66  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
; OTHER INFORMATION: Peptide  
US-09-357-952-66

Query Match 61.5%; Score 32; DB 4; Length 10;  
Best Local Similarity 50.0%; Pred. No. 1.5;  
Matches 5; Conservative 3; Mismatches 2; Indels 2; Gaps 0;

Qy 1 EEVVPXGMSY 10  
::: || |  
Db 1 DDIVPCSMY 10

RESULT 9  
US-09-521-650-66  
; Sequence 66, Application US/09521650  
; Patent No. 6335429  
; GENERAL INFORMATION:  
; APPLICANT: Weber, Eckard

; APPLICANT: Cai, Sui Xiong  
 ; APPLICANT: Keana, John F.W.  
 ; APPLICANT: Drewe, John A.  
 ; APPLICANT: Zhang, Han-Zhong  
 ; TITLE OF INVENTION: No. 6335429el Fluorogenic or Fluorescent Reporter Molecules and  
 ; TITLE OF INVENTION: Their Applications for Whole-Cell Fluorescence  
 ; TITLE OF INVENTION: Screening Assays for Caspases and Other Enzymes and the  
 ; TITLE OF INVENTION: Use Thereof  
 ; FILE REFERENCE: 1735.0290002  
 ; CURRENT APPLICATION NUMBER: US/09/521,650  
 ; CURRENT FILING DATE: 2000-03-08  
 ; EARLIER APPLICATION NUMBER: 09/168,888  
 ; EARLIER FILING DATE: 1998-10-09  
 ; EARLIER APPLICATION NUMBER: US 60/061,582  
 ; EARLIER FILING DATE: 1997-10-10  
 ; EARLIER APPLICATION NUMBER: US 09/033,661  
 ; EARLIER FILING DATE: 1998-03-03  
 ; NUMBER OF SEQ ID NOS: 142  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 66  
 ; LENGTH: 10  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Description of Artificial Sequence:Synthetic  
 ; OTHER INFORMATION: Peptide  
 US-09-521-650-66

Query Match 61.5%; Score 32; DB 4; Length 10;  
 Best Local Similarity 50.0%; Pred. No. 1.5;  
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EEVVPXGMSY 10  
 Db 1 DDIVPCMSY 10

RESULT 10  
 US-09-168-888-66  
 ; Sequence 66, Application US/09168888  
 ; Patent No. 6342611  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Weber, Eckard  
 ; APPLICANT: Cai, Sui Xiong  
 ; APPLICANT: Keana, John F.W.  
 ; APPLICANT: Drewe, John A.  
 ; APPLICANT: Zhang, Han-Zhong  
 ; TITLE OF INVENTION: No. 6342611el Fluorogenic or Fluorescent Reporter Molecules and  
 ; TITLE OF INVENTION: Their Applications for Whole-Cell Fluorescence  
 ; TITLE OF INVENTION: Screening Assays for Caspases and Other Enzymes and the  
 ; TITLE OF INVENTION: Use Thereof  
 ; FILE REFERENCE: 1735.0290002  
 ; CURRENT APPLICATION NUMBER: US/09/168,888  
 ; CURRENT FILING DATE: 1998-10-09  
 ; EARLIER APPLICATION NUMBER: US 60/061,582  
 ; EARLIER FILING DATE: 1997-10-10  
 ; EARLIER APPLICATION NUMBER: US 09/033,661  
 ; EARLIER FILING DATE: 1998-03-03  
 ; NUMBER OF SEQ ID NOS: 142  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 66  
 ; LENGTH: 10  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Description of Artificial Sequence:Synthetic  
 ; OTHER INFORMATION: Peptide  
 US-09-168-888-66

Query Match 61.5%; Score 32; DB 4; Length 10;  
 Best Local Similarity 50.0%; Pred. No. 1.5;  
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 EEVVPXGMSY 10  
 Db 1 DDIVPCMSY 10

## RESULT 11

US-08-580-988A-23  
 ; Sequence 23, Application US/08580988A  
 ; Patent No. 5856161  
 ; GENERAL INFORMATION:

; APPLICANT: Aggarwal et al.  
 ; TITLE OF INVENTION: Tumor Necrosis Factor  
 ; TITLE OF INVENTION: Receptor-I-Associated Protein Kinase And Methods  
 ; TITLE OF INVENTION: For Its Use  
 ; NUMBER OF SEQUENCES: 27  
 ; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Dr. Benjamin A. Adler  
 ; STREET: 8011 Candle Lane  
 ; CITY: Houston  
 ; STATE: Texas  
 ; COUNTRY: USA  
 ; ZIP: 77071

; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: 1.44 Mb floppy disk  
 ; COMPUTER: Apple Macintosh  
 ; OPERATING SYSTEM: Macintosh  
 ; SOFTWARE: Microsoft Word for Macintosh  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/580,988A  
 ; FILING DATE: January 3, 1996  
 ; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER:  
 ; FILING DATE:

; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Benjamin Aaron Adler, Ph.D., J.D.  
 ; REGISTRATION NUMBER: 35,423

; REFERENCE/DOCKET NUMBER: D5721CIP2  
 ; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 713-777-2321  
 ; TELEFAX: 713-777-6908

; INFORMATION FOR SEQ ID NO: 23:

; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 102 amino acids

; TYPE: amino acid

; STRANDEDNESS:

; TOPOLOGY: linear

; MOLECULE TYPE:

; DESCRIPTION: protein

; HYPOTHETICAL: no

; ANTI-SENSE: no

; FRAGMENT TYPE: internal

; ORIGINAL SOURCE:

US-08-580-988A-23

Query Match 61.5%; Score 32; DB 2; Length 102;  
 Best Local Similarity 60.0%; Pred. No. 21;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 EEVVPXGMSY 10  
 Db 24 EEVFPPLANNY 33

## RESULT 12

US-08-460-694-4  
 ; Sequence 4, Application US/08460694  
 ; Patent No. 5858655

; GENERAL INFORMATION:

; APPLICANT: Arnold, Andrew

; TITLE OF INVENTION: PRAD1 Cyclin and its cDNA

; NUMBER OF SEQUENCES: 8

; CORRESPONDENCE ADDRESS:



ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
STREET: 1100 New York Avenue, N.W., Suite 600  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/460,694  
FILING DATE: 02-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: McConathy, Evelyn H.  
REGISTRATION NUMBER: 35,279  
REFERENCE/DOCKET NUMBER: 0609.4070002  
TELEPHONE: 202-371-2600  
TELEFAX: 202-371-2540  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 152 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-460-694-4

Query Match 61.5%; Score 32; DB 2; Length 152;  
Best Local Similarity 60.0%; Pred. No. 33;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 EEVVPXGMSY 10  
||| | | |  
Db 20 EEVFPPLAMNY 29

RESULT 13  
US-08-460-744-4  
Sequence 4, Application US/08460744  
Patent No. 6107541  
GENERAL INFORMATION:  
APPLICANT: Arnold, Andrew  
TITLE OF INVENTION: PRADI Cyclin and its cDNA  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
STREET: 1100 New York Avenue, N.W., Suite 600  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/460,744  
FILING DATE: 02-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: McConathy, Evelyn H.  
REGISTRATION NUMBER: 35,279  
REFERENCE/DOCKET NUMBER: 0609.4070005  
TELEPHONE: 202-371-2600  
TELEFAX: 202-371-2540  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:

LENGTH: 152 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-460-744-4

Query Match 61.5%; Score 32; DB 3; Length 152;  
Best Local Similarity 60.0%; Pred. No. 33;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 EEVVPXGMSY 10  
||| | | |  
Db 20 EEVFPPLAMNY 29

RESULT 14  
US-07-667-711B-4  
Sequence 4, Application US/07667711B  
Patent No. 6110700  
GENERAL INFORMATION:  
APPLICANT: ARNOLD, ANDREW  
TITLE OF INVENTION: Pradi Cyclin and its cDNA  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
STREET: 1100 NEW YORK AVE., NW, SUITE 600  
CITY: WASHINGTON  
STATE: DC  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/667,711B  
FILING DATE: 11-MAR-1991  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: MCPHAIL, DONALD R.  
REGISTRATION NUMBER: 35,811  
REFERENCE/DOCKET NUMBER: 0609.4070000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 371-2600  
TELEFAX: (202) 371-2540  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 152 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: not relevant  
MOLECULE TYPE: peptide  
US-07-667-711B-4

Query Match 61.5%; Score 32; DB 3; Length 152;  
Best Local Similarity 60.0%; Pred. No. 33;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 EEVVPXGMSY 10  
||| | | |  
Db 20 EEVFPPLAMNY 29

RESULT 15  
US-08-193-977-7  
Sequence 7, Application US/08193977  
Patent No. 5625031  
GENERAL INFORMATION:  
APPLICANT: WEBSTER, KEVIN R.  
APPLICANT: COLEMAN, KEVIN G.  
TITLE OF INVENTION: PEPTIDE INHIBITORS OF THE P33CDK2 AND

;; TITLE OF INVENTION: P34CDC2 CELL CYCLE REGULATORY KINASES AND HUMAN  
;; NUMBER OF SEQUENCES: 34  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: REED & ROBINS  
;; STREET: 635 BRYANT STREET  
;; CITY: PALO ALTO  
;; STATE: CALIFORNIA  
;; COUNTRY: UNITED STATES OF AMERICA  
;; ZIP: 94301  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/193,977  
;; FILING DATE: 08-FEB-1994  
;; CLASSIFICATION: 530  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: ROBINS, ROBERTA L.  
;; REGISTRATION NUMBER: 33,208  
;; REFERENCE/DOCKET NUMBER: 5998-0016  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (415) 617-8999  
;; TELEFAX: (415) 327-3231  
;; INFORMATION FOR SEQ ID NO: 7:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 173 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
;; US-08-193-977-7

Query Match 61.5%; Score 32; DB 1; Length 173;  
Best Local Similarity 60.0%; Pred. No. 38;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EVVVPXGMSY 10  
||| | | |  
DB 55 EEVFPPLAMNY 64

RESULT 16  
US-08-464-517-21  
;; Sequence 21, Application US/08464517  
;; Patent No. 5869640  
;; GENERAL INFORMATION:  
;; APPLICANT: BEACH, David H.  
;; TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO  
;; NUMBER OF SEQUENCES: 50  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: LAHIVE & COCKFIELD  
;; STREET: 60 State Street  
;; CITY: Boston  
;; STATE: MA  
;; COUNTRY: USA  
;; ZIP: 02109  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: ASCII(text)  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/464,517  
;; FILING DATE:  
;; CLASSIFICATION: 435  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/963,308  
;; FILING DATE: 16-OCT-1992  
;; APPLICATION NUMBER: US 07/888,178  
;; FILING DATE: 26-MAY-1992

;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/701,514  
;; FILING DATE: 16-MAY-1991  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Matthew P. Vincent  
;; REGISTRATION NUMBER: 36,709  
;; REFERENCE/DOCKET NUMBER: MII-004C  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (617) 227-7400  
;; TELEFAX: (617) 227-5941  
;; INFORMATION FOR SEQ ID NO: 21:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 189 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
;; US-08-464-517-21

Query Match 61.5%; Score 32; DB 2; Length 189;  
Best Local Similarity 60.0%; Pred. No. 42;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EVVVPXGMSY 10  
||| | | |  
DB 74 EEVFPPLAMNY 83

RESULT 17  
US-08-246-361A-21  
;; Sequence 21, Application US/08246361A  
;; Patent No. 5998582  
;; GENERAL INFORMATION:  
;; APPLICANT: BEACH, David H.  
;; TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO  
;; NUMBER OF SEQUENCES: 50  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: LAHIVE & COCKFIELD  
;; STREET: 60 State Street  
;; CITY: Boston  
;; STATE: MA  
;; COUNTRY: USA  
;; ZIP: 02109  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: ASCII(text)  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/246,361A  
;; FILING DATE: 19-MAY-1994  
;; CLASSIFICATION: 435  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/963,308  
;; FILING DATE: 16-OCT-1992  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/888,178  
;; FILING DATE: 26-MAY-1992  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/701,514  
;; FILING DATE: 16-MAY-1991  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Matthew P. Vincent  
;; REGISTRATION NUMBER: 36,709  
;; REFERENCE/DOCKET NUMBER: MII-004C  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (617) 227-7400  
;; TELEFAX: (617) 227-5941  
;; INFORMATION FOR SEQ ID NO: 21:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 189 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single

TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-246-361A-21

Query Match 61.5%; Score 32; DB 2; Length 189;  
Best Local Similarity 60.0%; Pred. No. 42;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 EEVVPXGMSY 10  
||| | | :  
Db 74 EEVFPPLAMNY 83

## RESULT 18

US-08-463-772-21  
Sequence 21, Application US/08463772  
Patent No. 6066501

GENERAL INFORMATION:  
APPLICANT: BEACH, David H.  
TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO  
NUMBER OF SEQUENCES: 50  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD  
STREET: 60 State Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII(text)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/463,772  
FILING DATE:  
CLASSIFICATION: 435

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/963,308  
FILING DATE: 16-OCT-1992  
APPLICATION NUMBER: US 07/888,178  
FILING DATE: 26-MAY-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/701,514  
FILING DATE: 16-MAY-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Matthew P. Vincent  
REGISTRATION NUMBER: 36,709  
REFERENCE/DOCKET NUMBER: MII-004C  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 227-7400  
TELEFAX: (617) 227-5941

INFORMATION FOR SEQ ID NO: 21:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 189 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-463-772-21

Query Match 61.5%; Score 32; DB 3; Length 189;  
Best Local Similarity 60.0%; Pred. No. 42;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 EEVVPXGMSY 10  
||| | | :  
Db 74 EEVFPPLAMNY 83

## RESULT 19

PCT-US93-05000-21  
Sequence 21, Application PC/TUS9305000

GENERAL INFORMATION:  
APPLICANT: MITOTIX  
TITLE OF INVENTION: D-Type Cyclin and Uses Related Thereto  
NUMBER OF SEQUENCES: 42  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
STREET: Two Militia Drive  
CITY: Lexington  
STATE: Massachusetts  
COUNTRY: US  
ZIP: 02173  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US93/05000  
FILING DATE: 19930525  
CLASSIFICATION:

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/888,178  
FILING DATE: 26-MAY-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Granahan, Patricia  
REGISTRATION NUMBER: 32,227  
REFERENCE/DOCKET NUMBER: CSHL91-02A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-861-6240  
TELEFAX: 616-861-9540  
INFORMATION FOR SEQ ID NO: 21:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 189 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
PCT-US93-05000-21

Query Match 61.5%; Score 32; DB 5; Length 189;  
Best Local Similarity 60.0%; Pred. No. 42;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 EEVVPXGMSY 10  
||| | | :  
Db 74 EEVFPPLAMNY 83

## RESULT 20

US-08-464-517-22  
Sequence 22, Application US/08464517  
Patent No. 5869640

GENERAL INFORMATION:  
APPLICANT: BEACH, David H.  
TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO  
NUMBER OF SEQUENCES: 50  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD  
STREET: 60 State Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII(text)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/464,517  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/963,308

;; FILING DATE: 16-OCT-1992  
;; APPLICATION NUMBER: US 07/888,178  
;; FILING DATE: 26-MAY-1992  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/701,514  
;; FILING DATE: 16-MAY-1991  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Matthew P. Vincent  
;; REGISTRATION NUMBER: 36,709  
;; REFERENCE/DOCKET NUMBER: MII-004C  
;; TELEPHONE: (617) 227-7400  
;; TELEFAX: (617) 227-5941  
;; INFORMATION FOR SEQ ID NO: 22:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 236 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
US-08-464-517-22

Query Match 61.5%; Score 32; DB 2; Length 236;  
Best Local Similarity 60.0%; Pred. No. 54;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10  
||| | | |  
Db 20 EEVFLPMNY 29

## RESULT 21

US-08-246-361A-22  
;; Sequence 22, Application US/08246361A  
;; Patent No. 5998582  
;; GENERAL INFORMATION:  
;; APPLICANT: BEACH, David H.  
;; TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO  
;; NUMBER OF SEQUENCES: 50  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: LAHIVE & COCKFIELD  
;; STREET: 60 State Street  
;; CITY: Boston  
;; STATE: MA  
;; COUNTRY: USA  
;; ZIP: 02109  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: ASCII(text)  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/246,361A  
;; FILING DATE: 19-MAY-1994  
;; CLASSIFICATION: 435  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/963,308  
;; FILING DATE: 16-OCT-1992  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/888,178  
;; FILING DATE: 26-MAY-1992  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/701,514  
;; FILING DATE: 16-MAY-1991  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Matthew P. Vincent  
;; REGISTRATION NUMBER: 36,709  
;; REFERENCE/DOCKET NUMBER: MII-004C  
;; TELEPHONE: (617) 227-7400  
;; TELEFAX: (617) 227-5941  
;; INFORMATION FOR SEQ ID NO: 22:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 236 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
US-08-464-517-22

;; LENGTH: 236 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
US-08-246-361A-22

Query Match 61.5%; Score 32; DB 2; Length 236;  
Best Local Similarity 60.0%; Pred. No. 54;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10  
||| | | |  
Db 20 EEVFLPMNY 29

## RESULT 22

US-08-463-772-22  
;; Sequence 22, Application US/08463772  
;; Patent No. 6066501  
;; GENERAL INFORMATION:  
;; APPLICANT: BEACH, David H.  
;; TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO  
;; NUMBER OF SEQUENCES: 50  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: LAHIVE & COCKFIELD  
;; STREET: 60 State Street  
;; CITY: Boston  
;; STATE: MA  
;; COUNTRY: USA  
;; ZIP: 02109  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: ASCII(text)  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/463,772  
;; FILING DATE:  
;; CLASSIFICATION: 435  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/963,308  
;; FILING DATE: 16-OCT-1992  
;; APPLICATION NUMBER: US 07/888,178  
;; FILING DATE: 26-MAY-1992  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/701,514  
;; FILING DATE: 16-MAY-1991  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Matthew P. Vincent  
;; REGISTRATION NUMBER: 36,709  
;; REFERENCE/DOCKET NUMBER: MII-004C  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (617) 227-7400  
;; TELEFAX: (617) 227-5941  
;; INFORMATION FOR SEQ ID NO: 22:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 236 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
US-08-463-772-22

Query Match 61.5%; Score 32; DB 3; Length 236;  
Best Local Similarity 60.0%; Pred. No. 54;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10  
||| | | |  
Db 20 EEVFLPMNY 29

RESULT 23  
PCT-US93-05000-22  
; Sequence 22, Application PC/TUS9305000  
; GENERAL INFORMATION:  
; APPLICANT: MITOTIX  
; TITLE OF INVENTION: D-Type Cyclin and Uses Related Thereto  
; NUMBER OF SEQUENCES: 42  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
; STREET: Two Militia Drive  
; CITY: Lexington  
; STATE: Massachusetts  
; COUNTRY: US  
; ZIP: 02173  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US93/05000  
; FILING DATE: 19930525  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/07/888,178  
; FILING DATE: 26-MAY-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Granahan, Patricia  
; REGISTRATION NUMBER: 32,227  
; REFERENCE/DOCKET NUMBER: CSHL91-02A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-861-6240  
; TELEFAX: 616-861-9540  
; INFORMATION FOR SEQ ID NO: 22:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 236 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
PCT-US93-05000-22

Query Match 61.5%; Score 32; DB 5; Length 236;  
Best Local Similarity 60.0%; Pred. No. 54;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 BEVPXGMSY 10  
| | | | |  
Db 20 BEVFLPMNY 29

RESULT 24  
US-08-464-517-6  
; Sequence 6, Application US/08464517  
; Patent No. 5869640  
; GENERAL INFORMATION:  
; APPLICANT: BEACH, David H.  
; TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO  
; NUMBER OF SEQUENCES: 50  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD  
; STREET: 60 State Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII(text)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/464,517  
; FILING DATE:

; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/963,308  
; FILING DATE: 16-OCT-1992  
; APPLICATION NUMBER: US 07/888,178  
; FILING DATE: 26-MAY-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/701,514  
; FILING DATE: 16-MAY-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Matthew P. Vincent  
; REGISTRATION NUMBER: 36,709  
; REFERENCE/DOCKET NUMBER: MII-004C  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 227-7400  
; TELEFAX: (617) 227-5941  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 280 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-464-517-6

Query Match 61.5%; Score 32; DB 2; Length 280;  
Best Local Similarity 60.0%; Pred. No. 65;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 BEVPXGMSY 10  
| | | | |  
Db 75 BEVFLPMNY 84

RESULT 25  
US-08-463-772-6  
; Sequence 6, Application US/08463772  
; Patent No. 6066501  
; GENERAL INFORMATION:  
; APPLICANT: BEACH, David H.  
; TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO  
; NUMBER OF SEQUENCES: 50  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD  
; STREET: 60 State Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII(text)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/463,772  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/963,308  
; FILING DATE: 16-OCT-1992  
; APPLICATION NUMBER: US 07/888,178  
; FILING DATE: 26-MAY-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/701,514  
; FILING DATE: 16-MAY-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Matthew P. Vincent  
; REGISTRATION NUMBER: 36,709  
; REFERENCE/DOCKET NUMBER: MII-004C  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 227-7400  
; TELEFAX: (617) 227-5941  
; INFORMATION FOR SEQ ID NO: 6:

; SEQUENCE CHARACTERISTICS:  
; LENGTH: 280 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-463-772-6

Query Match 61.5%; Score 32; DB 3; Length 280;  
Best Local Similarity 60.0%; Pred. NO. 65;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10  
||| | | |  
Db 75 EEVFPPLAMNY 84

Search completed: June 10, 2003, 13:51:33  
Job time : 10.6429 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2003, 13:46:50 ; Search time 15 Seconds  
(without alignments)  
75.710 Million cell updates/sec

Title: US-09-909-164-9  
Perfect score: 52  
Sequence: 1 EEEVXPXGMSYS 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 392085 seqs, 103240269 residues

Total number of hits satisfying chosen parameters: 392085

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 75 summaries

Database : Published\_Applications\_AA:\*  
1: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB pep.\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB pep.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB pep.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB pep.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB pep.\*  
6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB pep.\*  
7: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB pep.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB pep.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB pep.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB pep.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB pep.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB pep.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB pep.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	38	73.1	3472	9	US-10-027-806-4
2	38	73.1	3472	9	US-10-034-623-4
3	38	73.1	3472	9	US-10-027-801-4
4	34	65.4	947	9	US-10-101-464A-73
5	33	63.5	426	9	US-10-214-766-43
6	33	63.5	478	9	US-09-924-340-108
7	33	63.5	478	9	US-09-992-600A-108
8	33	63.5	478	9	US-09-746-783-184
9	33	63.5	478	9	US-10-000-489-108
10	33	63.5	478	9	US-10-000-986-108
11	33	63.5	653	9	US-09-820-843A-26
12	32	61.5	10	10	US-09-947-387-66
13	32	61.5	254	10	US-09-778-927A-53
14	32	61.5	289	9	US-10-024-066-2
15	32	61.5	289	9	US-10-024-066-4
16	32	61.5	289	10	US-09-919-457-54
17	32	61.5	295	10	US-09-925-300-1061
18	32	61.5	529	10	US-09-923-304-4
19	32	61.5	691	9	US-10-101-921-4

Sequence 2, Appli	691	10	US-09-925-731-2
Sequence 10384, A	1377	9	US-09-815-242-10384
Sequence 4, Appli	2799	9	US-10-151-736-4
Sequence 878, App	53	9	US-10-032-154-878
Sequence 878, App	53	10	US-09-764-847-878
Sequence 14, Appl	53	10	US-09-948-080-14
Sequence 5124, Ap	161	9	US-09-738-626-5124
Sequence 20, Appl	163	9	US-10-117-846-20
Sequence 171, App	192	9	US-09-986-480-171
Sequence 334, App	198	10	US-09-731-872-334
Sequence 6349, App	223	9	US-09-738-626-6349
Sequence 10697, A	223	10	US-09-815-242-10697
Sequence 5, Appli	381	9	US-09-975-139-5
Sequence 2, Appli	702	9	US-10-280-403-2
Sequence 2, Appli	702	10	US-09-907-479-2
Sequence 4454, Ap	763	9	US-09-738-626-4454
Sequence 5136, Ap	1053	10	US-09-815-242-5136
Sequence 10439, A	1407	10	US-09-815-242-10439
Sequence 340, App	1426	10	US-09-912-020-340
Sequence 1, Appli	7	9	US-09-909-062-1
Sequence 9, Appli	7	9	US-09-909-062-9
Sequence 130, App	7	9	US-09-909-062-130
Sequence 68, Appl	121	9	US-09-852-797-68
Sequence 85, Appl	121	9	US-09-852-797-85
Sequence 68, Appl	121	10	US-09-853-161-68
Sequence 85, Appl	121	10	US-09-853-161-85
Sequence 68, Appl	121	10	US-09-852-659A-68
Sequence 85, Appl	121	10	US-09-852-659A-85
Sequence 359, App	135	9	US-09-992-598-359
Sequence 359, App	135	9	US-09-989-293A-359
Sequence 359, App	135	9	US-09-989-735-359
Sequence 359, App	135	9	US-09-990-444-359
Sequence 359, App	135	9	US-09-989-730-359
Sequence 359, App	135	9	US-09-990-436-359
Sequence 359, App	135	9	US-09-991-181-359
Sequence 359, App	135	9	US-09-983-687-359
Sequence 359, App	135	9	US-09-989-734-359
Sequence 359, App	135	9	US-09-997-653-359
Sequence 444, App	135	9	US-10-174-590-444
Sequence 444, App	135	9	US-10-176-758-444
Sequence 444, App	135	9	US-10-175-737-444
Sequence 444, App	135	9	US-09-993-667-359
Sequence 444, App	135	9	US-10-173-706-444
Sequence 444, App	135	9	US-10-175-738-444
Sequence 444, App	135	9	US-10-175-752-444
Sequence 444, App	135	9	US-10-176-482-444
Sequence 444, App	135	9	US-10-176-757-444
Sequence 444, App	135	9	US-10-176-913-444
Sequence 444, App	135	9	US-10-180-552-444
Sequence 444, App	135	9	US-10-180-557-444
Sequence 359, App	135	9	US-09-990-438-359
Sequence 359, App	135	9	US-09-990-562-359
Sequence 359, App	135	9	US-09-997-428-359
Sequence 359, App	135	9	US-09-997-666-359
Sequence 444, App	135	9	US-10-173-700-444
Sequence 444, App	135	9	US-10-174-572-444

## ALIGNMENTS

## RESULT 1

US-10-027-806-4  
; Sequence 4, Application US/10027806  
; Patent No. US20020160476A1  
; GENERAL INFORMATION:  
; APPLICANT: Swanson, Robert V.  
; APPLICANT: Feldman, Robert A.  
; APPLICANT: Schleper, Christa  
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM CENARCHAUM SYMBIOSUM  
; FILE REFERENCE: DORP.002A  
; CURRENT APPLICATION NUMBER: US/10/027,806  
; CURRENT FILING DATE: 2001-12-21

; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/408,020  
 ; PRIOR FILING DATE: EARLIER FILING DATE: 1999-09-29  
 ; NUMBER OF SEQ ID NOS: 123  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 4  
 ; LENGTH: 3472  
 ; TYPE: PRT  
 ; ORGANISM: Cenarchaeum symbiosum  
 US-10-027-806-4

Query Match 73.1%; Score 38; DB 9; Length 3472;  
 Best Local Similarity 54.5%; Pred. No. 1.2e+02;  
 Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXGMSYS 11  
 |:|:| |:|:|  
 Db 2294 EDVIPRGISFS 2304

RESULT 2  
 US-10-034-623-4  
 ; Sequence 4, Application US/10034623  
 ; Publication No. US20020198365A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Swanson, Ronald V.  
 ; APPLICANT: Feldman, Robert A.  
 ; APPLICANT: Schleper, Christa  
 ; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM CENARCHAEUM SYMBIOSUM  
 ; FILE REFERENCE: DCCP.002A  
 ; CURRENT APPLICATION NUMBER: US/10/034,623  
 ; CURRENT FILING DATE: 2001-12-21  
 ; PRIOR APPLICATION NUMBER: 09/408,020  
 ; PRIOR FILING DATE: 1999-09-29  
 ; PRIOR APPLICATION NUMBER: 60/102,294  
 ; PRIOR FILING DATE: 1998-09-29  
 ; NUMBER OF SEQ ID NOS: 123  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 4  
 ; LENGTH: 3472  
 ; TYPE: PRT  
 ; ORGANISM: Cenarchaeum symbiosum  
 US-10-034-623-4

Query Match 73.1%; Score 38; DB 9; Length 3472;  
 Best Local Similarity 54.5%; Pred. No. 1.2e+02;  
 Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXGMSYS 11  
 |:|:| |:|:|  
 Db 2294 EDVIPRGISFS 2304

RESULT 3  
 US-10-027-801-4  
 ; Sequence 4, Application US/10027801  
 ; Publication No. US20030054364A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Swanson, Ronald V.  
 ; APPLICANT: Feldman, Robert A.  
 ; APPLICANT: Schleper, Christa  
 ; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM CENARCHAEUM SYMBIOSUM  
 ; FILE REFERENCE: DCCP.002A  
 ; CURRENT APPLICATION NUMBER: US/10/027,801  
 ; CURRENT FILING DATE: 2001-12-21  
 ; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/408,020  
 ; PRIOR FILING DATE: EARLIER FILING DATE: 1999-09-29  
 ; NUMBER OF SEQ ID NOS: 123  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 4  
 ; LENGTH: 3472  
 ; TYPE: PRT  
 ; ORGANISM: Cenarchaeum symbiosum  
 US-10-027-801-4

Query Match 73.1%; Score 38; DB 9; Length 3472;  
 Best Local Similarity 54.5%; Pred. No. 1.2e+02;  
 Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXGMSYS 11  
 |:|:| |:|:|  
 Db 2294 EDVIPRGISFS 2304

RESULT 4  
 US-10-101-464A-73  
 ; Sequence 73, Application US/10101464A  
 ; Publication No. US20030046728A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Strabala, Timothy  
 ; APPLICANT: Nieuwenhuizen, Nicolaas  
 ; APPLICANT: Higgins, Colleen M.  
 ; TITLE OF INVENTION: Compositions Isolated from Plant Cells  
 ; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signaling  
 ; FILE REFERENCE: 11000.1020c2  
 ; CURRENT APPLICATION NUMBER: US/10/101,464A  
 ; CURRENT FILING DATE: 2002-03-18  
 ; PRIOR APPLICATION NUMBER: 09/704,302  
 ; PRIOR FILING DATE: 2000-11-01  
 ; PRIOR APPLICATION NUMBER: 09/228,986  
 ; PRIOR FILING DATE: 1999-01-12  
 ; PRIOR APPLICATION NUMBER: 60/162,866  
 ; PRIOR FILING DATE: 1999-11-01  
 ; PRIOR APPLICATION NUMBER: PCT/US00/00724  
 ; PRIOR FILING DATE: 2000-01-11  
 ; NUMBER OF SEQ ID NOS: 989  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 73  
 ; LENGTH: 947  
 ; TYPE: PRT  
 ; ORGANISM: Pinus radiata  
 US-10-101-464A-73

Query Match 65.4%; Score 34; DB 9; Length 947;  
 Best Local Similarity 66.7%; Pred. No. 1.8e+02;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMSYS 11  
 |:|:| |:|:|  
 Db 686 VMPGSGISYS 694

RESULT 5  
 US-10-214-766-43  
 ; Sequence 43, Application US/10214766  
 ; Publication No. US20030084473A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gocal, Greg  
 ; TITLE OF INVENTION: NON-TRANSGENIC HERBICIDE RESISTANT PLANTS  
 ; FILE REFERENCE: CA1138  
 ; CURRENT APPLICATION NUMBER: US/10/214,766  
 ; CURRENT FILING DATE: 2002-08-09  
 ; PRIOR APPLICATION NUMBER: US 60/311,734  
 ; PRIOR FILING DATE: 2001-08-09  
 ; NUMBER OF SEQ ID NOS: 53  
 ; SOFTWARE: Patent in version 3.2  
 ; SEQ ID NO 43  
 ; LENGTH: 426  
 ; TYPE: PRT  
 ; ORGANISM: Vibrio cholerae  
 US-10-214-766-43

Query Match 63.5%; Score 33; DB 9; Length 426;  
 Best Local Similarity 60.0%; Pred. No. 1.2e+02;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10



Wed Jun 11 15:45:05 2003

Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 EVVPXGMSYS 11  
|||  
Db 239 EVAPAGASYN 248

RESULT 6

US-09-924-340-108  
; Sequence 108, Application US/09924340  
; Publication No. US20030027248A1  
; GENERAL INFORMATION:  
; APPLICANT: Benjamin, Stephanie  
; APPLICANT: Tanaka, Hiroaki  
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF  
; FILE REFERENCE: 91.052.REG  
; CURRENT APPLICATION NUMBER: US/09/924,340  
; CURRENT FILING DATE: 2001-08-06  
; PRIOR APPLICATION NUMBER: US 60/305,456  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: US 60/302,277  
; PRIOR FILING DATE: 2001-06-29  
; PRIOR APPLICATION NUMBER: US 60/298,698  
; PRIOR FILING DATE: 2001-06-15  
; PRIOR APPLICATION NUMBER: US 60/293,574  
; PRIOR FILING DATE: 2001-05-25  
; NUMBER OF SEQ ID NOS: 112  
; SOFTWARE: Jpatent  
; SEQ ID NO 108  
; LENGTH: 478  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-924-340-108

Query Match 63.5%; Score 33; DB 9; Length 478;  
Best Local Similarity 60.0%; Pred. No. 1.4e+02;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 EVVPXGMSYS 11  
|||  
Db 239 EVAPAGASYN 248

RESULT 7

US-09-992-600A-108  
; Sequence 108, Application US/0992600A  
; Publication No. US20030027161A1  
; GENERAL INFORMATION:  
; APPLICANT: Benjamin, Stephanie  
; APPLICANT: Tanaka, Hiroaki  
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF  
; FILE REFERENCE: 91.054.DIV  
; CURRENT APPLICATION NUMBER: US/09/992,600A  
; CURRENT FILING DATE: 2001-11-13  
; PRIOR APPLICATION NUMBER: US 09/924,340  
; PRIOR FILING DATE: 2001-08-06  
; PRIOR APPLICATION NUMBER: PCT/IB01/01715  
; PRIOR FILING DATE: 2001-08-06  
; PRIOR APPLICATION NUMBER: US 60/305,456  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: US 60/302,277  
; PRIOR FILING DATE: 2001-06-29  
; PRIOR APPLICATION NUMBER: US 60/298,698  
; PRIOR FILING DATE: 2001-06-15  
; PRIOR APPLICATION NUMBER: US 60/293,574  
; PRIOR FILING DATE: 2001-05-25  
; NUMBER OF SEQ ID NOS: 114  
; SOFTWARE: Jpatent  
; SEQ ID NO 108  
; LENGTH: 478  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-992-600A-108

Query Match 63.5%; Score 33; DB 9; Length 478;  
Best Local Similarity 60.0%; Pred. No. 1.4e+02;

Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 EVVPXGMSYS 11  
|||  
Db 239 EVAPAGASYN 248

RESULT 8

US-09-746-783-184  
; Sequence 184, Application US/09746783  
; Publication No. US20030044935A1  
; GENERAL INFORMATION:  
; APPLICANT: Jacobs, Kenneth  
; APPLICANT: McCoy, John M.  
; APPLICANT: Lavallie, Edward R.  
; APPLICANT: Racie, Lisa A.  
; APPLICANT: Treacy, Maurice  
; APPLICANT: Spaulding, Vikki  
; APPLICANT: Agostino, Michael J.  
; APPLICANT: Howes, Steven H.  
; APPLICANT: Fechtel, Kim  
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES  
; NUMBER OF SEQUENCES: 231  
; CORRESPONDENCE ADDRESS:  
; STREET: 87 Cambridgepark Drive  
; CITY: Cambridge  
; STATE: MA  
; COUNTRY: U.S.A.  
; ZIP: 02140  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/746,783  
; FILING DATE: 21-Dec-2000  
; CLASSIFICATION: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Milasincic, Debra J.  
; REGISTRATION NUMBER: 46,931  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 227-7400  
; TELEFAX: (617) 742-4214  
; INFORMATION FOR SEQ ID NO: 184:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 478 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: <Unknown>  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 184:  
US-09-746-783-184

Query Match 63.5%; Score 33; DB 9; Length 478;  
Best Local Similarity 60.0%; Pred. No. 1.4e+02;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 EVVPXGMSYS 11  
|||  
Db 239 EVAPAGASYN 248

RESULT 9

US-10-000-489-108  
; Sequence 108, Application US/10000489  
; Publication No. US20030092011A1  
; GENERAL INFORMATION:  
; APPLICANT: Benjamin, Stephanie  
; APPLICANT: Tanaka, Hiroaki  
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF  
; FILE REFERENCE: 91.052.REG  
; CURRENT APPLICATION NUMBER: US/10/000,489  
; CURRENT FILING DATE: 2001-11-13  
; PRIOR APPLICATION NUMBER: US 10/024,340  
; PRIOR FILING DATE: 2001-08-06  
; PRIOR APPLICATION NUMBER: PCT/IB01/01715  
; PRIOR FILING DATE: 2001-08-06  
; PRIOR APPLICATION NUMBER: US 60/305,456  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: US 60/302,277  
; PRIOR FILING DATE: 2001-06-29  
; PRIOR APPLICATION NUMBER: US 60/298,698  
; PRIOR FILING DATE: 2001-06-15  
; PRIOR APPLICATION NUMBER: US 60/293,574  
; PRIOR FILING DATE: 2001-05-25  
; NUMBER OF SEQ ID NOS: 114  
; SOFTWARE: Jpatent  
; SEQ ID NO 108  
; LENGTH: 478  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-000-489-108

Query Match 63.5%; Score 33; DB 9; Length 478;  
Best Local Similarity 60.0%; Pred. No. 1.4e+02;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

;; FILE REFERENCE: 91.US6.DIV  
;; CURRENT APPLICATION NUMBER: US/10/000,489  
;; PRIOR FILING DATE: 2001-11-14  
;; PRIOR APPLICATION NUMBER: US 09/924,340  
;; PRIOR FILING DATE: 2001-08-06  
;; PRIOR APPLICATION NUMBER: PCT/IB01/01715  
;; PRIOR FILING DATE: 2001-08-06  
;; PRIOR APPLICATION NUMBER: US 60/305,456  
;; PRIOR FILING DATE: 2001-07-13  
;; PRIOR APPLICATION NUMBER: US 60/302,277  
;; PRIOR FILING DATE: 2001-06-29  
;; PRIOR APPLICATION NUMBER: US 60/298,698  
;; PRIOR FILING DATE: 2001-06-15  
;; PRIOR APPLICATION NUMBER: US 60/293,574  
;; PRIOR FILING DATE: 2001-05-25  
;; NUMBER OF SEQ ID NOS: 112  
;; SOFTWARE: JPatent  
;; SEQ ID NO 108  
;; LENGTH: 478  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-10-000-489-108

Query Match 63.5%; Score 33; DB 9; Length 478;  
Best Local Similarity 60.0%; Pred. No. 1.4e+02;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 EVVPXGMSYS 11  
||| |||  
Db 239 EVAPAGASYN 248

## RESULT 10

US-10-000-986-108  
;; Sequence 108, Application US/10000986  
;; Publication No. US20030096247A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Benjamin, Stephanie  
;; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF  
;; FILE REFERENCE: 91.US9.DIV  
;; CURRENT APPLICATION NUMBER: US/10/000,986  
;; PRIOR FILING DATE: 2001-11-14  
;; PRIOR APPLICATION NUMBER: US 09/924,340  
;; PRIOR FILING DATE: 2001-08-06  
;; PRIOR APPLICATION NUMBER: PCT/IB01/01715  
;; PRIOR FILING DATE: 2001-08-06  
;; PRIOR APPLICATION NUMBER: US 60/305,456  
;; PRIOR FILING DATE: 2001-07-13  
;; PRIOR APPLICATION NUMBER: US 60/302,277  
;; PRIOR FILING DATE: 2001-06-29  
;; PRIOR APPLICATION NUMBER: US 60/298,698  
;; PRIOR FILING DATE: 2001-06-15  
;; PRIOR APPLICATION NUMBER: US 60/293,574  
;; PRIOR FILING DATE: 2001-05-25  
;; NUMBER OF SEQ ID NOS: 112  
;; SOFTWARE: JPatent  
;; SEQ ID NO 108  
;; LENGTH: 478  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-10-000-986-108

Query Match 63.5%; Score 33; DB 9; Length 478;  
Best Local Similarity 60.0%; Pred. No. 1.4e+02;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 EVVPXGMSYS 11  
||| |||  
Db 239 EVAPAGASYN 248

## RESULT 11

US-09-820-843A-26  
;; Sequence 26, Application US/09820843A  
;; Publication No. US20030039963A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Council of Scientific and Industrial Research  
;; TITLE OF INVENTION: A COMPUTATIONAL METHOD FOR THE IDENTIFICATION OF CANDIDATE  
;; FILE REFERENCE: Q63915  
;; CURRENT APPLICATION NUMBER: US/09/820,843A  
;; PRIOR FILING DATE: 2001-03-30  
;; NUMBER OF SEQ ID NOS: 118  
;; SOFTWARE: PatentIn version 3.0  
;; SEQ ID NO 26  
;; LENGTH: 653  
;; TYPE: PRT  
;; ORGANISM: Vibrio cholerae  
;; NAME/KEY: misc-feature  
;; OTHER INFORMATION: Iron(III) ABC transporter, permease protein  
;; NAME/KEY: misc-feature  
;; OTHER INFORMATION: gi|9654609  
US-09-820-843A-26

Query Match 63.5%; Score 33; DB 9; Length 653;  
Best Local Similarity 66.7%; Pred. No. 1.9e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXGMS 9  
||||| |||  
Db 300 EEVVPXGMS 308

## RESULT 12

US-09-947-387-66  
;; Sequence 66, Application US/09947387  
;; Patent No. US20020150885A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Weber, Eckard  
;; APPLICANT: Cai, Sui Xiong  
;; APPLICANT: Keana, John F.W.  
;; APPLICANT: Drewe, John A.  
;; APPLICANT: Zhang, Han-Zhong  
;; TITLE OF INVENTION: No. US20020150885A1 Fluorogenic or Fluorescent Reporter Mo  
;; TITLE OF INVENTION: Their Applications for Whole-Cell Fluorescence  
;; TITLE OF INVENTION: Screening Assays for Caspases and Other Enzymes and the  
;; TITLE OF INVENTION: Use Thereof  
;; FILE REFERENCE: 1735.0290005  
;; CURRENT APPLICATION NUMBER: US/09/947,387  
;; PRIOR FILING DATE: 2001-09-07  
;; PRIOR APPLICATION NUMBER: US 60/061,582  
;; PRIOR FILING DATE: 1997-10-10  
;; PRIOR APPLICATION NUMBER: US 60/145,746  
;; PRIOR FILING DATE: 1998-03-03  
;; PRIOR APPLICATION NUMBER: US 09/168,888  
;; PRIOR FILING DATE: 1998-10-09  
;; NUMBER OF SEQ ID NOS: 142  
;; SOFTWARE: PatentIn Ver. 2.0  
;; SEQ ID NO 66  
;; LENGTH: 10  
;; TYPE: PRT  
;; ORGANISM: Artificial Sequence  
;; FEATURE:  
;; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-09-947-387-66

Query Match 61.5%; Score 32; DB 10; Length 10;  
Best Local Similarity 50.0%; Pred. No. 3.5;  
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10  
::: ||| |||  
Db 1 DDIVPCMSY 10

us-09-909-164-9.rapb

Wed Jun 11 15:45:05 2003

```

; Sequence 4, Application US/10024066
; Patent No. US20020166134A1
; GENERAL INFORMATION:
; APPLICANT: Field, Loren J.
; APPLICANT: Pasumathi, Kishore Babu S.
; TITLE OF INVENTION: CARDIOMYOCYTES WITH ENHANCED PROLIFERATIVE POTENTIAL,
; TITLE OF INVENTION: AND METHODS FOR PREPARING AND USING SAME
; FILE REFERENCE: 7037-450
; CURRENT APPLICATION NUMBER: US/10/024,066
; CURRENT FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: 60/139,942
; PRIOR FILING DATE: 1999-06-18
; PRIOR APPLICATION NUMBER: PCT/US00/16827
; PRIOR FILING DATE: 2000-06-19
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 289
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-024-066-4

Query Match          61.5%  Score 32; DB 9; Length 289;
Best Local Similarity 60.0%  Pred. No. 1.3e+02;
Matches              6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY      1 EEVVPXGMSY 10
      ||| | | |
Db      74 EEVFPPLAMNY 83

RESULT 16
US-09-919-497-54
; Sequence 54, Application US/09919497
; Patent No. US20020106662A1
; GENERAL INFORMATION:
; APPLICANT: Mutter, George L.
; TITLE OF INVENTION: PROGNOSTIC CLASSIFICATION OF ENDOMETRIAL CANCER
; FILE REFERENCE: B0801/7225
; CURRENT APPLICATION NUMBER: US/09/919,497
; CURRENT FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: US 60/221,735
; PRIOR FILING DATE: 2000-07-31
; NUMBER OF SEQ ID NOS: 100
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 54
; LENGTH: 289
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-919-497-54

Query Match          61.5%  Score 32; DB 10; Length 289;
Best Local Similarity 60.0%  Pred. No. 1.3e+02;
Matches              6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY      1 EEVVPXGMSY 10
      ||| | | |
Db      74 EEVFPPLAMNY 83

RESULT 17
US-09-925-300-1061
; Sequence 1061, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben,
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08

; Sequence 53, Application US/09778927A
; Patent No. US20020068342A1
; GENERAL INFORMATION:
; APPLICANT: KHOSRAVI, Rami et al.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID AND AMINO ACID SEQUENCES AND NOVEL
; TITLE OF INVENTION: VARIANTS OF ALTERNATIVE SPLICING
; FILE REFERENCE: 2786-0160P
; CURRENT APPLICATION NUMBER: US/09/778,927A
; CURRENT FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: IL 134453
; PRIOR FILING DATE: 2000-02-09
; PRIOR APPLICATION NUMBER: IL135341
; PRIOR FILING DATE: 2000-03-29
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 53
; LENGTH: 254
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)..(254 )
; OTHER INFORMATION: Xaa = any amino acid, unknown, or other
; US-09-778-927A-53

Query Match          61.5%  Score 32; DB 10; Length 254;
Best Local Similarity 60.0%  Pred. No. 1.1e+02;
Matches              6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY      1 EEVVPXGMSY 10
      ||| | | |
Db      74 EEVFPPLAMNY 83

RESULT 14
US-10-024-066-2
; Sequence 2, Application US/10024066
; Patent No. US20020166134A1
; GENERAL INFORMATION:
; APPLICANT: Field, Loren J.
; APPLICANT: Pasumathi, Kishore Babu S.
; TITLE OF INVENTION: CARDIOMYOCYTES WITH ENHANCED PROLIFERATIVE POTENTIAL,
; TITLE OF INVENTION: AND METHODS FOR PREPARING AND USING SAME
; FILE REFERENCE: 7037-450
; CURRENT APPLICATION NUMBER: US/10/024,066
; CURRENT FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: 60/139,942
; PRIOR FILING DATE: 1999-06-18
; PRIOR APPLICATION NUMBER: PCT/US00/16827
; PRIOR FILING DATE: 2000-06-19
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 289
; TYPE: PRT
; ORGANISM: Mus musculus
; US-10-024-066-2

Query Match          61.5%  Score 32; DB 9; Length 289;
Best Local Similarity 60.0%  Pred. No. 1.3e+02;
Matches              6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY      1 EEVVPXGMSY 10
      ||| | | |
Db      73 EEVFPPLAMNY 82

RESULT 15
US-10-024-066-4
```

```
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1061
; LENGTH: 295
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (243)
; OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (277)
; OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids
US-09-925-300-1061

Query Match
Best Local Similarity 61.5%; Score 32; DB 10; Length 295;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 EVVPXGMSYS 11
Db 52 EVLPKMSYA 61

RESULT 18
US-09-923-304-4
; Sequence 4, Application US/09923304
; Patent No. US20020081612A1
; GENERAL INFORMATION:
; APPLICANT: KATZ, RUTH
; APPLICANT: JIANG, FENG
; TITLE OF INVENTION: DETECTION AND DIAGNOSIS OF SMOKING RELATED CANCERS
; FILE REFERENCE: UTSC:658US
; CURRENT APPLICATION NUMBER: US/09/923.304
; CURRENT FILING DATE: 2001-08-06
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 529
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-923-304-4

Query Match
Best Local Similarity 61.5%; Score 32; DB 10; Length 529;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EVVPXGMSY 10
Db 48 QEVLPGLKY 57

RESULT 19
US-10-101-921-4
; Sequence 4, Application US/10101921
; Publication No. US20030022199A1
; GENERAL INFORMATION:
; APPLICANT: Nezu, Jun-ichi
; APPLICANT: Ose, Asuka
; APPLICANT: Tsuji, Akira
; TITLE OF INVENTION: TRANSPORTER GENES OATP-B, C, D, AND E
; FILE REFERENCE: 06501-104US1
; CURRENT APPLICATION NUMBER: US/10/101,921
; CURRENT FILING DATE: 2002-06-28
; PRIOR APPLICATION NUMBER: PCT/JP00/06416
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: JP 11/267835
; PRIOR FILING DATE: 1999-09-21
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
```

```
; LENGTH: 691
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-101-921-4

Query Match
Best Local Similarity 61.5%; Score 32; DB 9; Length 691;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMSY 10
Db 188 IVPLGLSY 195

RESULT 20
US-09-925-731-2
; Sequence 2, Application US/09925731
; Patent No. US20020090622A1
; GENERAL INFORMATION:
; APPLICANT: ADEKUN, ANTHONI MONISOLA
; APPLICANT: AMERSE, HELEN JEAN
; APPLICANT: CRESSWELL, CARL JOHN
; APPLICANT: DUDLEY, ADAM JESTON
; TITLE OF INVENTION: CHEMICAL COMPOUNDS
; FILE REFERENCE: DJB/009901/0282795
; CURRENT APPLICATION NUMBER: US/09/925.731
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/226,909
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 691
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-731-2

Query Match
Best Local Similarity 61.5%; Score 32; DB 10; Length 691;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMSY 10
Db 188 IVPLGLSY 195

RESULT 21
US-09-815-242-10384
; Sequence 10384, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
```

NAME/KEY: misc\_feature  
LOCATION: (7)  
OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids  
US-10-092-154-878

Query Match 59.6%; Score 31; DB 9; Length 53;  
Best Local Similarity 66.7%; Pred. No. 33;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 VVPXGMSYS 11  
| | | : | | |  
Db 39 VVPTAVSYS 47

RESULT 24  
US-09-764-847-878  
Sequence 878, Application US/09764847  
Patent No. US20020132767A1  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
FILE REFERENCE: PC009  
CURRENT APPLICATION NUMBER: US/09/764,847  
CURRENT FILING DATE: 2001-01-17  
Prior application data removed - consult PALM or file wrapper  
NUMBER OF SEQ ID NOS: 2003  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 878  
LENGTH: 53  
TYPE: PRT  
ORGANISM: Homo sapiens  
NAME/KEY: SITE  
LOCATION: (7)  
OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids  
US-09-764-847-878

Query Match 59.6%; Score 31; DB 10; Length 53;  
Best Local Similarity 66.7%; Pred. No. 33;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 VVPXGMSYS 11  
| | | : | | |  
Db 39 VVPTAVSYS 47

RESULT 25  
US-09-948-080-14  
Sequence 14, Application US/09948080  
Patent No. US20020102702A1  
GENERAL INFORMATION:  
APPLICANT: VAN DER OSTEN, CLAU  
APPLICANT: HALKIER, TORDEN  
APPLICANT: ANDERSEN, CARSTEN  
APPLICANT: BAUDITZ, PETER  
APPLICANT: HANSEN, PETER KAMP  
TITLE OF INVENTION: PROTEASE VARIANTS AND COMPOSITIONS  
FILE REFERENCE: 4946,200-US  
CURRENT APPLICATION NUMBER: US/09/948,080  
CURRENT FILING DATE: 2001-09-06  
PRIOR APPLICATION NUMBER: US/08/963,851  
PRIOR FILING DATE: 1997-11-04  
NUMBER OF SEQ ID NOS: 35  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 14  
LENGTH: 59  
TYPE: PRT  
ORGANISM: Enterococcus faecalis  
US-09-948-080-14

Query Match 59.6%; Score 31; DB 10; Length 59;  
Best Local Similarity 45.5%; Pred. No. 37;  
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

PRIOR FILING DATE: 2000-11-27  
PRIOR APPLICATION NUMBER: 60/257,931  
PRIOR FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: 60/269,308  
PRIOR FILING DATE: 2001-02-16  
NUMBER OF SEQ ID NOS: 14110  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 10384  
LENGTH: 1377  
TYPE: PRT  
ORGANISM: Escherichia coli  
US-09-815-242-10384

Query Match 61.5%; Score 32; DB 10; Length 1377;  
Best Local Similarity 45.5%; Pred. No. 6.7e+02;  
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMSYS 11  
| : | | | | |  
Db 369 EQLNPAGLSYT 379

RESULT 22  
US-10-151-736-4  
Sequence 4, Application US/10151736  
Publication No. US20020192160A1  
GENERAL INFORMATION:  
APPLICANT: Callaghan, Michelle J.  
APPLICANT: Sutherland, Lindfield  
APPLICANT: Watts, Colin K.  
TITLE OF INVENTION: NO. US20020192160A1 Human Tumour Suppressor Gene  
FILE REFERENCE: RICE-010CON  
CURRENT APPLICATION NUMBER: US/10/151,736  
CURRENT FILING DATE: 2002-05-15  
PRIOR APPLICATION NUMBER: 09/403,402  
PRIOR FILING DATE: 1999-10-21  
PRIOR APPLICATION NUMBER: PC9/AU98/00280  
PRIOR FILING DATE: 1998-04-20  
NUMBER OF SEQ ID NOS: 17  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 4  
LENGTH: 2799  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-151-736-4

Query Match 61.5%; Score 32; DB 9; Length 2799;  
Best Local Similarity 60.0%; Pred. No. 1.4e+03;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 EVVVPXGMSYS 11  
| | : | | | |  
Db 2096 EVLPTKMSYA 2105

RESULT 23  
US-10-092-154-878  
Sequence 878, Application US/10092154  
Publication No. US20030054375A1  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
FILE REFERENCE: PC009C1  
CURRENT APPLICATION NUMBER: US/10/092,154  
CURRENT FILING DATE: 2002-03-07  
NUMBER OF SEQ ID NOS: 2003  
Prior Application removed - See File Wrapper or Palm  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 878  
LENGTH: 53  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:

Wed Jun 11 15:45:05 2003

us-09-909-164-9.rapb

Page 8

QY 1 EEWFXGMSYS 11  
|: :| |: |  
Db 38 EKHPGGLEYS 48

Search completed: June 10, 2003, 14:35:42  
Job time : 15.0714 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2003, 13:31:15 ; Search time 11.2143 Seconds  
(without alignments)  
94.297 Million cell updates/sec

Title: US-09-909-164-9  
Perfect score: 52  
Sequence: 1 EEVVPXGMSYS 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues  
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 75 summaries

Database : PIR.73:\*  
1: Pirl:\*  
2: Pirl2:\*  
3: Pirl3:\*  
4: Pirl4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	38	73.1	3472	2 T31308	hypothetical 367K
2	37	71.2	840	2 T39116	probable sulfate p
3	37	71.2	877	2 T40413	sulfate permease -
4	36	69.2	102	2 A42452	V1 protein - tobac
5	36	69.2	1498	2 B97355	DNA segregation At
6	35	67.3	225	2 S57810	hypothetical prote
7	35	67.3	425	2 T24111	hypothetical prote
8	35	67.3	670	2 S22293	zinc finger protei
9	35	67.3	749	2 H82691	topoisomerase IV s
10	35	67.3	2717	2 A34203	DNA-binding protei
11	34	65.4	156	2 S54619	hypothetical prote
12	34	65.4	252	2 H69491	cell division inhi
13	34	65.4	544	2 C92900	probable ABC subst
14	33	63.5	94	2 I40758	hypothetical prote
15	33	63.5	116	2 E90544	50S ribosomal prot
16	33	63.5	165	2 D69493	hypothetical prote
17	33	63.5	253	2 C81374	hypothetical prote
18	33	63.5	259	2 T34536	hypothetical prote
19	33	63.5	284	2 S75817	hypothetical prote
20	33	63.5	298	2 T47670	beta-ketoacyl-ACP
21	33	63.5	368	2 F72281	hypothetical prote
22	33	63.5	426	2 D82163	3-phosphoshikimate
23	33	63.5	466	2 T43653	cdc37 protein - fi
24	33	63.5	653	2 D82352	iron(III) ABC tran
25	33	63.5	890	2 A30481	bacteriocin BGN5 -
26	33	63.5	1028	2 AF3286	ATP-dependent DNA
27	33	63.5	1152	2 D87046	conserved hypothet
28	33	63.5	1394	2 A35626	transforming growt
29	33	63.5	1401	2 G82336	DNA-directed RNA p

30	63.5	1548	2 T04456	hypothetical prote
31	63.5	1712	2 A38261	masking protein pr
32	61.5	84	2 E97333	hypothetical prote
33	61.5	175	2 P00616	transport protein
34	61.5	223	2 T01457	rho protein Gbp-di
35	61.5	279	2 B72481	hypothetical prote
36	61.5	288	2 JC4011	cyclin D2 - rat
37	61.5	288	2 I58372	cyclin D2 - mouse
38	61.5	289	2 A41984	cyclin D2 - human
39	61.5	289	2 A42822	cyclin D2 - Africa
40	61.5	291	2 S57922	cyclin D2 - Africa
41	61.5	291	2 S57925	cyclin D2 - chicke
42	61.5	291	2 JC4579	cyclin D1 - zebra
43	61.5	291	2 S62730	cyclin D3 - human
44	61.5	292	2 B42822	cyclin D1 - human
45	61.5	295	2 A38977	cyclin D1 - mouse
46	61.5	295	2 A35523	cyclin D1 - rat
47	61.5	295	2 JC2342	cyclin D1 - mouse
48	61.5	347	2 I55120	rhsf protein - Esc
49	61.5	363	2 D69551	conserved hypothet
50	61.5	427	2 F64064	tolB protein - Hae
51	61.5	449	2 A92866	conserved hypothet
52	61.5	498	2 B90604	conserved hypothet
53	61.5	525	2 D98311	agaA protein limpo
54	61.5	525	2 AF2971	hypothetical prote
55	61.5	726	2 T44000	transport protein
56	61.5	726	2 T44187	infected cell prot
57	61.5	759	2 S25330	SCrl protein - yea
58	61.5	889	2 S22659	hypothetical prote
59	61.5	922	2 AG1827	maltoiligosyltreha
60	61.5	993	1 P1VXTA	RNA 1 protein - to
61	61.5	1306	2 S22624	aggregation protei
62	61.5	1377	2 C65159	rhsA protein precu
63	61.5	1377	2 E86034	rhsA protein in rh
64	61.5	1394	2 H91236	RhsH core protein
65	61.5	1397	2 A85570	rhsC protein in rh
66	61.5	1397	2 C64805	RhsC core protein
67	61.5	1399	2 A99720	RhsA core protein
68	61.5	1409	2 F91187	RhsB protein precu
69	61.5	1411	2 E65145	gene 11-1 protein
70	61.5	1948	2 S00485	trans-regulatory s
71	59.6	124	1 VKLJST	hypothetical prote
72	59.6	133	2 A71173	probable cobH - My
73	59.6	208	1 D70764	probable purQ prot
74	59.6	224	2 G70709	phosphoribosylform
75	59.6	224	2 F87186	

ALIGNMENTS

RESULT 1

T31308 hypothetical 367K protein - Cenarchaeum symbiosum

C:Species: Cenarchaeum symbiosum

C:Date: 11-Jan-2000 #sequence\_revision 11-Jan-2000 #text\_change 18-Feb-2000

C:Accession: T31308

R:Schleper, C.; DeLong, E.F.; Preston, C.M.; Feldman, R.A.; Wu, K.Y.; Swanson, R.V

J. Bacteriol. 180, 5003-5009, 1998

A:Title: Genomic analysis reveals chromosomal variation in natural populations of

A:Reference number: Z20994; MUID:98422450; PMID:9748430

A:Accession: T31308

A>Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-3472 <SC>

A:Cross-references: EMBL:AF083072; NID:g3599393; PID:g3599394; PIDN:AAC62699.1

C:Superfamily: Cenarchaeum symbiosum hypothetical 367K protein

Query Match 73.1% Score 38; DB 2; Length 3472;  
Best Local Similarity 54.5%; Pred. No. 59;  
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXGMSYS 11

Db 2294 EDVIPRGISFS 2304

## RESULT 2

T39116  
probable sulfate permease - fission yeast (Schizosaccharomyces pombe)  
C:Species: Schizosaccharomyces pombe  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Dec-1999  
C:Accession: T39116  
R:Hunt, C.; Aves, S.; McDougall, R.C.; Rajandream, M.A.; Barrell, B.G.  
submitted to the EMBL Data Library, November 1999  
A:Reference number: Z21829  
A:Accession: T39116  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-840 <HUN>  
A:Cross-references: EMBL:AL132779; PIDN:CAB60015.1; GSPDB:GN000066; SPDB:SPAC869.05c  
A:Experimental source: strain 972h-; cosmid c869  
C:Genetics:  
A:Gene: SPDB:SPAC869.05c  
A:Map position: 1

Query Match 71.2%; Score 37; DB 2; Length 840;  
Best Local Similarity 77.8%; Pred. No. 21;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMSYS 11  
||| |||||  
Db 135 VVPOGMSYA 143

## RESULT 3

T40413  
sulfate permease - fission yeast (Schizosaccharomyces pombe)  
C:Species: Schizosaccharomyces pombe  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Dec-1999  
C:Accession: T40413  
R:Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Jimenez Martinez, J.  
submitted to the EMBL Data Library, August 1998  
A:Reference number: Z21926  
A:Accession: T40413  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-877 <LYN>  
A:Cross-references: EMBL:AL031261; PIDN:CAA20298.1; GSPDB:GN000067; SPDB:SPBC3H7.02  
A:Experimental source: strain 972h-; cosmid c3H7  
C:Genetics:  
A:Gene: SPDB:SPBC3H7.02  
A:Map position: 2

Query Match 71.2%; Score 37; DB 2; Length 877;  
Best Local Similarity 77.8%; Pred. No. 22;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMSYS 11  
||| |||||  
Db 148 VVPOGMSYA 156

## RESULT 4

A42452  
V1 protein - tobacco yellow dwarf virus (strain Australia)  
C:Species: tobacco yellow dwarf virus  
C:Date: 15-Jan-1993 #sequence\_revision 15-Jan-1993 #text\_change 08-Oct-1999  
C:Accession: A42452  
R:Morris, B.A.M.; Richardson, K.A.; Haley, A.; Zhan, X.; Thomas, J.E.  
Virology 187, 633-642, 1992  
A:Title: The nucleotide sequence of the infectious cloned DNA component of tobacco yellow  
A:Reference number: A42452; MUID:92188538; PMID:1546458  
A:Accession: A42452  
A:Molecule type: DNA  
A:Residues: 1-102 <MOR>

A:Cross-references: GB:M81103; NID:g335283; PIDN:AAA47947.1; PID:g335284  
Query Match 69.2%; Score 36; DB 2; Length 102;  
Best Local Similarity 60.0%; Pred. No. 3.5;  
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 EVVVPXGMSYS 11  
:||| |:|:  
Db 7 QVVPNGINS 16

## RESULT 5

B97355  
DNA segregation ATPase, FtsK/SpoIIIE family, YUKA B. subtilis ortholog [imported]  
C:Species: Clostridium acetobutylicum  
C:Date: 14-Sep-2001 #sequence\_revision 14-Sep-2001 #text\_change 14-Sep-2001  
C:Accession: B97355  
R:Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.;  
Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.  
J. Bacteriol. 183, 4823-4838, 2001  
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium  
A:Reference number: A96900; MUID:21359325; PMID:21359325  
A:Accession: B97355  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1498 <KUR>  
A:Cross-references: GB:AE001437; PIDN:AAK81629.1; PID:g15026814; GSPDB:GN00168  
A:Experimental source: Clostridium acetobutylicum ATCC824  
C:Genetics:  
A:Gene: CAC3709

Query Match 69.2%; Score 36; DB 2; Length 1498;  
Best Local Similarity 60.0%; Pred. No. 63;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10  
|:|:| |||||  
Db 1276 EQKIPMGMSY 1285

## RESULT 6

S57810  
hypothetical protein precursor (clone TP11) - tomato  
C:Species: Lycopersicon esculentum (tomato)  
C:Date: 28-Oct-1995 #sequence\_revision 03-Nov-1995 #text\_change 21-Jul-2000  
C:Accession: S57810  
R:Milligan, S.B.; Gasser, C.S.  
Plant Mol. Biol. 28, 691-711, 1995  
A:Title: Nature and regulation of pistil-expressed genes in tomato.  
A:Reference number: S57808; MUID:95375233; PMID:7647301  
A:Accession: S57810  
A:Status: preliminary; nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1-225 <MTL>  
A:Cross-references: EMBL:U20592; NID:g924625; PIDN:AAA80497.1; PID:g924626  
C:Superfamily: Plant Kunitz-type proteinase inhibitor

Query Match 67.3%; Score 35; DB 2; Length 225;  
Best Local Similarity 54.5%; Pred. No. 13;  
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMSYS 11  
:||| |:|:  
Db 32 DEVVPGKTYA 42

## RESULT 7

T24111  
hypothetical protein R10D12.10 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
C:Accession: T24111  
R:Percy, C.



submitted to the EMBL Data Library, October 1996

A:Reference number: Z19842  
 A:Accession: T24111  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-425 <WIL>  
 A:Cross-references: EMBL:Z81109; PIDN:CAB03241.1; GSPDB:GN00023; CESP:R10D12.10  
 A:Experimental source: clone R10D12  
 C:Genetics:  
 A:Gene: CESP:R10D12.10  
 A:Map position: 5  
 A:Introns: 23/3; 56/3; 113/3; 257/2

Query Match 67.3%; Score 35; DB 2; Length 425;  
 Best Local Similarity 50.0%; Pred. No. 26;  
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EVVPXGMSY 10  
 I:|:|:| I:  
 Db 335 EQIVPGGLQY 344

#### RESULT 8

S22293  
 zinc finger protein AT-BP2 - rat (fragment)  
 C:Species: Rattus norvegicus (Norway rat)  
 C:Date: 29-Jan-1998 #sequence\_revision 06-Feb-1998 #text\_change 20-Sep-1999  
 C:Accession: S22293; I78656  
 R:Mitohmore, C.; Traboni, C.; Cortese, R.  
 Nucleic Acids Res. 19, 141-147, 1991  
 A:Title: Isolation of two cDNAs encoding zinc finger proteins which bind to the alpha 1-  
 A:Reference number: 158280; MUID:91187610; PMID:1901405  
 A:Accession: S22293  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-670 <MIT>  
 A:Cross-references: EMBL:X54250; NID:g57519; PIDN:CAA38151.1; PID:g57520  
 A:Note: the authors did not translate the codon for residue 1  
 C:Superfamily: HIV-EP2 enhancer-binding protein  
 C:Keywords: DNA binding; transcription regulation; zinc finger

Query Match 67.3%; Score 35; DB 2; Length 670;  
 Best Local Similarity 66.7%; Pred. No. 43;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMSY 11  
 |:|:|:| I:  
 Db 376 VVPAGLTYS 384

#### RESULT 9

H82691  
 topoisomerase IV subunit XF1353 [imported] - Xylella fastidiosa (strain 9a5c)  
 C:Species: Xylella fastidiosa  
 C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Sep-2000  
 C:Accession: H82691  
 R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen  
 Nature 406, 151-157, 2000  
 A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.  
 A:Reference number: A82515; MUID:20365717; PMID:10910347  
 A:Note: for a complete list of authors see reference number A59328 below  
 A:Accession: H82691  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-749 <SIM>  
 A:Cross-references: GB:AE003967; GB:AE003849; NID:g9106347; PIDN:AAF84162.1; GSPDB:GN001  
 R:Experimental source: strain 9a5c  
 R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A  
 Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H  
 as-Neto, E.; Docena, C.; El-Borry, H.; Facincani, A.P.; Ferreira, A.J.S.  
 submitted to GenBank, June 2000  
 A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm  
 J.D.; Junqueira, M.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig

chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Marti  
 A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki,  
 F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri,  
 Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sa  
 A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da S  
 M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.  
 A:Reference number: A59328  
 A:Contents: annotation  
 C:Genetics:  
 A:Gene: XF1353  
 C:Superfamily: DNA topoisomerase (ATP-hydrolyzing) chain A; phage T4 DNA topoisomer

Query Match 67.3%; Score 35; DB 2; Length 749;  
 Best Local Similarity 77.8%; Pred. No. 48;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 EVVPXGMSY 10  
 |:|:|:| I:  
 Db 526 EVDPSGMSY 534

#### RESULT 10

A34203  
 DNA-binding protein PRDII-BF1 - human  
 N:Alternate names: major histocompatibility complex enhancer-binding protein 1  
 C:Species: Homo sapiens (man)  
 C:Date: 22-Jun-1990 #sequence\_revision 22-Jun-1990 #text\_change 20-Sep-1999  
 C:Accession: A34203; A34779  
 R:Fan, C.M.; Mantatis, T.  
 Genes Dev. 4, 29-42, 1990  
 A:Title: A DNA-binding protein containing two widely separated zinc finger motifs  
 A:Reference number: A34203; MUID:90169514; PMID:2106471  
 A:Accession: A34203  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-2717 <FAN>  
 A:Cross-references: EMBL:X51435; NID:g38017; PIDN:CAA35798.1; PID:g38018  
 R:Balwin Jr., A.S.; LeClair, K.P.; Singh, H.; Sharp, P.A.  
 Mol. Cell. Biol. 10, 1406-1414, 1990  
 A:Title: A large protein containing zinc finger domains binds to related sequence  
 A:Reference number: A34779; MUID:90205817; PMID:2108316  
 A:Accession: A34779  
 A:Status: preliminary; nucleic acid sequence not shown  
 A:Molecule type: mRNA  
 A:Residues: 801-1072, 'N', 1072-1168, 'K', 1170-1225, 'V', 1227-1434, 'N', 1436-1607, 'I', 1  
 A:Cross-references: GB:M32019  
 C:Superfamily: HIV-EP2 enhancer-binding protein  
 C:Keywords: DNA binding; transcription regulation; zinc finger

Query Match 67.3%; Score 35; DB 2; Length 2717;  
 Best Local Similarity 66.7%; Pred. No. 1.9e+02;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMSY 11  
 |:|:|:| I:  
 Db 2405 VVPAGLTYS 2413

#### RESULT 11

S54619  
 hypothetical protein YOR013w - yeast (Saccharomyces cerevisiae)  
 N:Alternate names: hypothetical protein O2612; hypothetical protein YOL303.3  
 C:Species: Saccharomyces cerevisiae  
 C:Date: 08-Jul-1995 #sequence\_revision 01-Sep-1995 #text\_change 19-Apr-2002  
 C:Accession: S54619; S66879  
 R:de Haan, M.; Maarse, A.C.; Grivell, L.A.  
 submitted to the EMBL Data Library, May 1995  
 A:Reference number: S54617  
 A:Accession: S54619  
 A:Molecule type: DNA  
 A:Residues: 1-156 <DE>  
 A:Cross-references: EMBL:X87331; NID:g1041652; PIDN:CAA60762.1; PID:g829123  
 R:de Haan, M.; Grivell, L.A.; Maarse, A.C.

submitted to the Protein Sequence Database, July 1996  
 A:Reference number: S66877

A:Accession: S66879

A:Molecule type: DNA

A:Residues: 1-156 <DEW>

A:Cross-references: EMBL:Z74920; NID:g1420109; PIDN:CAA99201.1; PID:g1420111; MIPS:YOR01

A:Experimental source: strain S288C

C:Genetics:

A:Cross-references: SGD:S0005539

A:Map position: 15R

C:Superfamily: hypothetical protein YOR013w

Query Match 65.4%; Score 34; DB 2; Length 156;  
 Best Local Similarity 66.7%; Pred. No. 14;

Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 EVVPXGMSY 10

Db 50 EVMLGMDY 58

RESULT 12

H69491

cell division inhibitor (mind-2) homolog - Archaeoglobus fulgidus

C:Species: Archaeoglobus fulgidus

C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 26-Aug-1999

C:Accession: H69491

R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson

.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.

Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.

Nature 390, 364-370, 1997

A:Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artlach, P.; Kaine, B.P.; Sykes, S.

A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archae

A:Reference number: A69250; MUID:198049343; PMID:9389475

A:Accession: H69491

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-252 <KLE>

A:Cross-references: GB:AE000970; GB:AE000782; NID:g2689293; PIDN:AA889318.1; PID:g264860

C:Superfamily: cell division inhibitor mind

Query Match 65.4%; Score 34; DB 2; Length 252;

Best Local Similarity 75.0%; Pred. No. 24;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 EVVPXGMS 9

Db 81 EVIPAGMS 88

RESULT 13

C82900

Probable ABC substrate-binding protein, iron U359 [imported] - Ureaplasma urealyticum

C:Species: Ureaplasma urealyticum

C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Sep-2000

C:Accession: C82900

R:Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Cassell, G.H.

submitted to GenBank, February 2000

A:Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a mir

A:Reference number: A82870

A:Accession: C82900

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-544 <GLA>

A:Cross-references: GB:AE002133; GB:AF222894; NID:g68999339; PIDN:AAF30768.1; GSPDB:GN001

A:Experimental source: serovar 1

C:Genetics:

A:Gene: ABCsbp-5; U359

A:Genetic code: SGC3

Query Match 65.4%; Score 34; DB 2; Length 544;

Best Local Similarity 70.0%; Pred. No. 55;

Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 EVVPXGMSY 10

Db 135 EVVPHYLSY 144

RESULT 14

I40758

hypothetical protein 1 - Campylobacter jejuni (fragment)

C:Species: Campylobacter jejuni

C:Date: 16-Aug-1996 #sequence\_revision 16-Aug-1996 #text\_change 08-Oct-1999

C:Accession: I40758; S47317

R:Hani, E.K.; Chan, V.L.

J. Bacteriol. 177, 2396-2402, 1995

A:Title: Expression and characterization of Campylobacter jejuni benzoylglutamine an

A:Reference number: I40758; MUID:95247673; PMID:7730270

A:Accession: I40758

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-94 <RES>

A:Cross-references: EMBL:Z36940; NID:g535805; PIDN:CAA85392.1; PID:g535806

Query Match 63.5%; Score 33; DB 2; Length 94;

Best Local Similarity 55.6%; Pred. No. 13;

Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 EVVPXGMSY 10

Db 26 DIFPSGMSY 34

RESULT 15

E90544

50S ribosomal protein L20 [imported] - Mycoplasma pulmonis (strain UAB CTIP)

C:Species: Mycoplasma pulmonis

C:Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 03-Aug-2001

C:Accession: E90544

R:Chambaud, I.; Heilig, R.; Ferris, S.; Barbe, V.; Samson, D.; Galisson, F.; Mosze

Nucleic Acids Res. 29, 2145-2153, 2001

A:Title: The complete genome sequence of the murine respiratory pathogen Mycoplasma

A:Reference number: A99512; MUID:21267165; PMID:11353084

A:Accession: E90544

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-116 <KUR>

A:Cross-references: GB:AL44566; PID:g14089674; PIDN:CAC13434.1; GSPDB:GN00153

A:Experimental source: strain UAB CTIP

C:Genetics:

A:Gene: MPU\_2610

A:Genetic code: SGC3

C:Superfamily: Escherichia coli ribosomal protein L20

Query Match 63.5%; Score 33; DB 2; Length 116;

Best Local Similarity 77.8%; Pred. No. 17;

Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 VVPXGMSYS 11

Db 68 VRPLGMSYS 76

RESULT 16

D69493

hypothetical protein AF1949 - Archaeoglobus fulgidus

C:Species: Archaeoglobus fulgidus

C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 04-Mar-2000

C:Accession: D69493

R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; I

.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness

Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.

Nature 390, 364-370, 1997

A:Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artlach, P.; Kaine, B.P.; Syke

Smith, H.O.; Woese, C.R.; Venter, J.C.  
 A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaea

A:Reference number: A69250; MUID:98049343; PMID:9389475

A:Accession: D69493

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-165 <KLE>

A:Cross-references: GB:AE000968; GB:AE000782; NID:g2689291; PIDN:AA89307.1; PID:g264859

C:Superfamily: Archaeoglobus fulgidus hypothetical protein AF1949

Query Match 63.5%; Score 33; DB 2; Length 165;

Best Local Similarity 60.0%; Pred. No. 25;

Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEWVPXGMSY 10

||| ||| |||

Db 60 EESIPDGASY 69

RESULT 17

C81374

hypothetical protein Cj0990c [imported] - Campylobacter jejuni (strain NCTC 11168)

C:Species: Campylobacter jejuni

C:Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 03-Jun-2002

C:Accession: C81374

R:Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chilling

C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVleet, A.; Whitehead, S.; Barrel

Nature 403, 665-668, 2000

A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hyp

A:Reference number: A81250; MUID:20150912; PMID:10688204

A:Accession: C81374

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-253 <PAR>

A:Cross-references: GB:AL139076; GB:AL111168; NID:g6968128; PIDN:CAB73246.1; PID:g696842

A:Experimental source: serotype O2, strain NCTC 11168

C:Genetics:

A:Gene: Cj0990c

Query Match 63.5%; Score 33; DB 2; Length 253;

Best Local Similarity 55.6%; Pred. No. 39;

Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 EWPVPXGMSY 10

::| |||||

Db 185 DIFPSGMSY 193

RESULT 18

T34536

hypothetical protein DKFZp434C031.1 - human (fragment)

C:Species: Homo sapiens (man)

C:Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 29-Oct-1999

C:Accession: T34536

R:Poustka, A.; Wellenreuther, R.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.

submitted to the Protein Sequence Database, October 1999

A:Reference number: Z21540

A:Accession: T34536

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-259 <POU>

A:Cross-references: EMBL:AL122063

A:Experimental source: adult testis; clone DKFZp434C031

C:Genetics:

A:Note: DKFZp434C031.1

Query Match

Best Local Similarity 63.5%; Score 33; DB 2; Length 259;

Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 EWPVPXGMSY 11

||| ||| |||

Db 22 EVAPAGASYN 31

RESULT 19

S75817

hypothetical protein slr1275 - *Synechocystis* sp. (strain PCC 6803)

C:Species: *Synechocystis* sp.

A:Variety: PCC 6803

C:Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 08-Oct-1999

C:Accession: S75817

R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima

O, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.;

DNA Res. 3, 109-136, 1996

A:Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechoc*

S.

A:Reference number: S74322; MUID:97061201; PMID:8905231

A:Accession: S75817

A>Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-284 <KAN>

A:Cross-references: EMBL:D90913; GB:AB001339; NID:g1653348; PIDN:BAA18276.1; PID:d

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 63.5%; Score 33; DB 2; Length 284;

Best Local Similarity 55.6%; Pred. No. 44;

Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMSYS 11

||| ||| |||

Db 208 VIPAGVSYT 216

RESULT 20

T47670

beta-ketoacyl-ACP reductase-like protein - *Arabidopsis thaliana*

N:Alternate names: protein T26112.190

C:Species: *Arabidopsis thaliana* (mouse-ear cross)

C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 18-Aug-2000

C:Accession: T47670

R:Monfort, A.; Casacuberta, E.; Puigdomenech, P.; Mewes, H.W.; Lemcke, K.; Mayer,

submitted to the Protein Sequence database, February 2000

A:Reference number: Z24471

A:Accession: T47670

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-298 <MON>

A:Cross-references: EMBL:AL132954

A:Experimental source: cultivar Columbia; BAC clone T26112

C:Genetics:

A:Map position: 3

A:Introns: 25/3

A:Note: T26112.190

C:Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology

Query Match

Best Local Similarity 63.5%; Score 33; DB 2; Length 298;

Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMSYS 11

||| ||| |||

Db 187 IVPGGLAYS 195

RESULT 21

F72281

hypothetical protein TM1216 - *Thermotoga maritima* (strain MSB8)

C:Species: *Thermotoga maritima*

C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 21-Jul-2000

C:Accession: F72281

R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.;

Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richards

C.M.

A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from gene

A:Reference number: A72200; MUID:99287316; PMID:10360571  
 A:Accession: F72281  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-368 <ARN>  
 A:Cross-references: GB:AE001778; GB:AE000512; NID:g4981757; PIDN:AAD36291.1; PID:g498176  
 A:Experimental source: strain MSB8  
 C:Genetics:  
 A:Gene: TMi216  
 C:Superfamily: NADH dehydrogenase (ubiquinone) 49K protein

Query Match 63.5%; Score 33; DB 2; Length 368;  
 Best Local Similarity 55.6%; Pred. No. 59;  
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMSYS 11  
 :|| |||:  
 Db 294 IVPGMAYA 302

## RESULT 22

D82163  
 3-phosphoshikimate 1-carboxyvinyltransferase VC1732 [imported] - Vibrio cholerae (strain  
 C:Species: Vibrio cholerae  
 C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Feb-2001  
 C:Accession: D82163  
 R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;  
 chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P.A.;  
 L, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
 Nature 406, 477-483, 2000  
 A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.  
 A:Reference number: A82035; MUID:20406833; PMID:10952301  
 A:Accession: D82163  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-426 <HEI>  
 A:Cross-references: GB:AE004251; GB:AE003852; NID:g9656248; PIDN:AAF94882.1; GSPDB:GN001  
 A:Experimental source: serogroup O1; strain N16961; biotype El Tor  
 C:Genetics:  
 A:Gene: VC1732  
 A:Map position: 1  
 C:Superfamily: 3-phosphoshikimate 1-carboxyvinyltransferase; 3-phosphoshikimate 1-carboxy

Query Match 63.5%; Score 33; DB 2; Length 426;  
 Best Local Similarity 60.0%; Pred. No. 69;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10  
 :|| |||:  
 Db 223 EFVIPAGQSY 232

## RESULT 23

T43653  
 cdc37 protein - fission yeast (Schizosaccharomyces pombe)  
 C:Species: Schizosaccharomyces pombe  
 C:Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 18-Feb-2000  
 C:Accession: T43653; T40791; T43654  
 R:Westwood, P.K.; Preston, N.C.; Fantes, P.A.  
 submitted to the EMBL Data Library, March 1999  
 A:Reference number: Z22602  
 A:Accession: T43653  
 A:Status: preliminary; translated from GB/EMBL/DBDJ  
 A:Molecule type: mRNA  
 A:Residues: 1-466 <WES>  
 A:Cross-references: EMBL:AJ132376; PIDN:CAB38757.1  
 R:Wood, V.; Rajandream, M.A.; Barrell, B.G.; Volckaert, G.  
 submitted to the EMBL Data Library, May 1999  
 A:Reference number: Z21875  
 A:Accession: T40791  
 A:Status: preliminary; translated from GB/EMBL/DBDJ  
 A:Molecule type: DNA

A:Residues: 1-466 <WOO>  
 A:Cross-references: EMBL:AL049769; PIDN:CAB42371.2; GSPDB:GN00067  
 A:Experimental source: strain 972h-; cosmid c9B6  
 R:Westwood, P.K.; Preston, N.C.; Fantes, P.A.  
 submitted to the EMBL Data Library, March 1999  
 A:Description: Schizosaccharomyces pombe cdc37 gene.  
 A:Reference number: Z22603  
 A:Accession: T43654  
 A:Status: preliminary; translated from GB/EMBL/DBDJ  
 A:Molecule type: DNA  
 A:Residues: 1-466 <WE2>  
 A:Cross-references: EMBL:AJ132377; PIDN:CAB38758.1  
 C:Genetics:  
 A:Gene: cdc37; SPAC9B6.10  
 A:Map position: 2  
 A:Introns: 8/2; 17/2; 21/1

Query Match 63.5%; Score 33; DB 2; Length 466;  
 Best Local Similarity 50.0%; Pred. No. 76;  
 Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10  
 :|| |||:  
 Db 98 DSAIFGMSY 107

## RESULT 24

D82352  
 iron(III) ABC transporter, permease protein VC0203 [imported] - Vibrio cholerae (strain  
 C:Species: Vibrio cholerae  
 C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Feb-2001  
 C:Accession: D82352  
 R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson,  
 chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers,  
 L, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
 Nature 406, 477-483, 2000  
 A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.  
 A:Reference number: A82035; MUID:20406833; PMID:10952301  
 A:Accession: D82352  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-653 <HEI>  
 A:Cross-references: GB:AE004110; GB:AE003852; NID:g9654600; PIDN:AAF93379.1; GSPDB:  
 A:Experimental source: serogroup O1; strain N16961; biotype El Tor  
 C:Genetics:  
 A:Gene: VC0203  
 A:Map position: 1

Query Match 63.5%; Score 33; DB 2; Length 653;  
 Best Local Similarity 66.7%; Pred. No. 11e+02;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXGMS 9  
 :|| |||:  
 Db 300 EEVVPXGMS 308

## RESULT 25

A30481  
 bacteriocin BCN5 - Clostridium perfringens plasmid pIP404  
 C:Species: Clostridium perfringens  
 C:Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 24-Nov-1999  
 C:Accession: A30481; S03779  
 R:Garnier, T.; Cole, S.T.  
 J. Bacteriol. 168, 1189-1196, 1986  
 A:Title: Characterization of a bacteriocinogenic plasmid from Clostridium perfringens  
 A:Reference number: J70354; MUID:87057020; PMID:2877971  
 A:Accession: A30481  
 A:Molecule type: DNA  
 A:Residues: 1-890 <GAR>  
 A:Cross-references: GB:M32882; GB:J03309; NID:g150738; PIDN:AAA98249.1; PID:g150739  
 C:Genetics:  
 A:Gene: bcn

A:Genome: plasmid  
 C:Superfamily: Clostridium perfringens plasmid pIP404 bacteriocin BCN5  
 C:Keywords: Bacteriocin

Query Match 63.58; Score 33; DB 2; Length 890;  
 Best Local Similarity 66.74; Pred. No. 1.5e+02;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 EVVPXGMSY 10  
 ||||| :!  
 Db 170 EVVPGGFTY 178

Search completed: June 10, 2003, 13:49:13  
 Job time : 11.2143 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2003, 13:25:04 ; Search time 4.5 seconds  
(without alignments)  
101.387 Million cell updates/sec

Title: US-09-909-164-9  
Perfect score: 52  
Sequence: 1 EEVVPXGMSYS 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 75 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	38	73.1	1058	1 CARB_FUSNN	Q8rg86 fuscobacteri
2	37	71.2	877	1 SULH_SCHPO	O74377 schizosacch
3	36	69.2	102	1 Y1LK_TYDVA	P31619 tobacco yel
4	36	69.2	1498	1 Y1A9_CLOAB	Q04351 clostridium
5	35	67.3	2717	1 ZEP1_HUMAN	P15822 homo sapien
6	34	65.4	788	1 C114_NEUCR	P23822 neosporea
7	34	65.4	1499	1 A10C_HUMAN	O60312 homo sapien
8	33	63.5	116	1 RL20_MYCPU	Q98q00 mycoplasma
9	33	63.5	253	1 Y990_CAMJE	P45489 campylobact
10	33	63.5	426	1 AROA_VIBCH	Q9krt0 vibrio chol
11	33	63.5	466	1 CC37_SCHPO	O94740 schizosacch
12	33	63.5	478	1 GSR2_HUMAN	Q9nzm5 homo sapien
13	33	63.5	890	1 BCN5_CLOPE	P08696 clostridium
14	33	63.5	1394	1 LTB5_HUMAN	P22064 homo sapien
15	33	63.5	1401	1 RPOC_VIBCH	Q9kv29 vibrio chol
16	33	63.5	1595	1 LTB1_HUMAN	Q14766 homo sapien
17	33	63.5	1712	1 LTB1_RAT	Q00918 rattus norv
18	32	61.5	288	1 CGD2_RAT	Q04827 rattus norv
19	32	61.5	289	1 CGD2_HUMAN	P30279 homo sapien
20	32	61.5	289	1 CGD2_MOUSE	P30280 mus musculu
21	32	61.5	291	1 CGD1_BRARE	Q90459 brachydanio
22	32	61.5	291	1 CGD1_XENLA	P50755 xenopus lae
23	32	61.5	291	1 CGD2_CHICK	P49706 gallus gall
24	32	61.5	291	1 CGD2_XENLA	P53782 xenopus lae
25	32	61.5	292	1 CGD1_CHICK	P55169 gallus gall
26	32	61.5	292	1 CGD3_HUMAN	P30281 homo sapien
27	32	61.5	295	1 CGD1_HUMAN	P24385 homo sapien
28	32	61.5	295	1 CGD1_MOUSE	P25322 mus musculu
29	32	61.5	295	1 CGD1_RAT	P39948 rattus norv
30	32	61.5	427	1 TOLB_HAEIN	P44677 haemophilus
31	32	61.5	529	1 ENP3_HUMAN	O75355 homo sapien
32	32	61.5	691	1 OAT6_HUMAN	Q9y616 homo sapien
33	32	61.5	726	1 PRTP_HSV6U	P52384 human herpe

34	61.5	759	1	SC11_YEAST	P32784 saccharomyc
35	61.5	920	1	EDD_RAT	Q62671 rattus norv
36	61.5	993	1	VIA_TAV	P28931 tomato aspe
37	61.5	1377	1	RHSA_ECOLI	P16916 escherichia
38	61.5	1397	1	RHSC_ECOLI	P16918 escherichia
39	61.5	1411	1	RHSB_ECOLI	P16917 escherichia
40	61.5	2799	1	EDD_HUMAN	O95071 homo sapien
41	61.5	124	1	REV_SIVCZ	P17280 chimpanzee
42	59.6	208	1	COBH_MYCTU	Q10676 mycobacteri
43	59.6	223	1	PURQ_PYRHO	O59619 pyrococcus
44	59.6	224	1	PURQ_HALNI	O9nnu2 halobacteri
45	59.6	224	1	PURQ_MYCLE	O05756 mycobacteri
46	59.6	224	1	PURQ_MYCTU	P71841 mycobacteri
47	59.6	225	1	PURQ_CORAM	Q9rhx0 corynebacte
48	59.6	240	1	GDIR_ARATH	O9sf66 arabidopsis
49	59.6	276	1	Y939_METJA	O58349 methanococc
50	59.6	319	1	YHAI_CRYPA	P10941 cryphonectr
51	59.6	432	1	PURA_YEAST	P80210 saccharomyc
52	59.6	488	1	NOM2_PSEAE	Q9htr0 pseudomonas
53	59.6	670	1	OATP_RAT	P46720 rattus norv
54	59.6	706	1	ADDG_HUMAN	Q9uey8 homo sapien
55	59.6	706	1	ADDC_MOUSE	Q9gyb5 mus musculu
56	59.6	827	1	PLSB_MOUSE	Q61586 mus musculu
57	59.6	828	1	PLSB_RAT	P97564 rattus norv
58	59.6	1047	1	EF3_SCHPO	O94489 schizosacch
59	59.6	1407	1	RPOC_ECOLI	P00577 escherichia
60	59.6	1426	1	RHSD_ECOLI	P16919 escherichia
61	59.6	2145	1	U520_CAEEL	Q9u2g0 caenorhabdi
62	58.7	472	1	ET2A_XENLA	P13102 xenopus lae
63	57.7	81	1	YB47_ARCFU	O28825 archaeroglob
64	57.7	121	1	TKNK_HUMAN	Q9uhf0 homo sapien
65	57.7	132	1	ATPE_AQUAE	Q6903 aquifex aeo
66	57.7	146	1	ATPE_LACAC	Q9tgy0 lactobacill
67	57.7	218	1	PURQ_METHH	O26370 methanobact
68	57.7	223	1	PURQ_RHINE	Q92pil rhizobium m
69	57.7	223	1	RP1A_BUCAL	P57489 buchnera ap
70	57.7	230	1	PURQ_METJA	O59042 methanococc
71	57.7	232	1	SCOA_HELFP	Q9zle3 helicobacte
72	57.7	232	1	SCOA_HELFP	P56006 helicobacte
73	57.7	286	1	CXA6_RAT	P82933 rattus norv
74	57.7	356	1	GBA2_USTMA	P87033 ustilago ma
75	57.7	361	1	RFB2_SALTIA	P26391 salmonella

ALIGNMENTS

RESULT 1				
CARB_FUSNN	STANDARD;	PRT;	1058 AA.	
ID	CARB_FUSNN	STANDARD;	PRT;	1058 AA.
AC	Q8RG86;			
DT	15-JUN-2002 (Rel. 41, Created)			
DT	15-JUN-2002 (Rel. 41, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Carbamoyl-phosphate synthase large chain (EC 6.3.5.5) (Carbamoyl-phosphate synthase ammonia chain).			
DE	phosphate synthase ammonia chain).			
GN	CARB OR FN0422			
OS	Fusobacterium nucleatum (subsp. nucleatum).			
OC	Bacteria; Fusobacteria; Fusobacterium.			
OX	NCBI_TaxID=76856;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=ATCC 25586;			
RX	MEDLINE=21886394; PubMed=11889109;			
RA	Kapatral V., Anderson I., Ivanova N., Resnik G., Los T., Lykidis A.,			
RA	Bhattacharyya A., Bartman A., Gardner W., Grechkin G., Zhu L.,			
RA	Vasileva O., Chu L., Kogan Y., Chaga O., Goltzman E., Bernal A.,			
RA	Larsen N., D'Souza M., Walunas T., Pusch G., Haselkorn R.,			
RA	Fonstein M., Kyripides N., Overbeek R.;			
RT	"Genome sequence and analysis of the oral bacterium Fusobacterium			
RT	nucleatum strain ATCC 25586."			
RL	J. Bacteriol. 184:2005-2018(2002).			
CC	-!- CATALYTIC ACTIVITY: 2 ATP + L-glutamine + CO(2) + H(2)O = 2 ADP +			

phosphate + L-glutamate + carbamoyl phosphate.  
 -1- COFACTOR: Binds three manganese ions (By similarity).  
 -1- PATHWAY: Arginine biosynthesis.  
 -1- SUBUNIT: Pyrimidine biosynthesis; first step.  
 -1- SUBUNIT: Composed of two chains; the small (or glutamine) chain promotes the hydrolysis of glutamine to ammonia, which is used by the large (or ammonia) chain to synthesize carbamoyl phosphate (By similarity).  
 -1- SIMILARITY: BELONGS TO THE CARB FAMILY.  
 -----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 EMBL; AF010554; AAL94625.1; ALT\_INIT.  
 InterPro; IPR005483; CPase\_L.  
 InterPro; IPR005479; CPase\_L\_D2.  
 InterPro; IPR005480; CPase\_L\_D3.  
 InterPro; IPR005481; CPase\_L\_N.  
 InterPro; IPR004362; MGS-like.  
 Pfam; PF00289; CPase\_L\_Chain; 2.  
 Pfam; PF02786; CPase\_L\_D2; 2.  
 Pfam; PF02787; CPase\_L\_D3; 1.  
 Pfam; PF02142; MGS; 1.  
 PRINTS; PR00098; CPASE.  
 PROSITE; PS00866; CPASE\_1; 2.  
 PROSITE; PS00867; CPASE\_2; 2.  
 Arginine biosynthesis; Pyrimidine biosynthesis; Ligase; Repeat;  
 KW ATP-binding; Manganese; Complete proteome.  
 FT DOMAIN 1 401 CARBOXYPHOSPHATE SYNTHETIC DOMAIN.  
 FT DOMAIN 402 546 OLIGOMERIZATION DOMAIN.  
 FT DOMAIN 547 929 CARBAMOYL PHOSPHATE SYNTHETIC DOMAIN.  
 FT DOMAIN 930 1058 ALLOSTERIC DOMAIN.  
 FT REPEAT 1 546  
 FT REPEAT 547 1058  
 FT NP\_BIND 153 210 ATP (POTENTIAL).  
 FT NP\_BIND 302 352 ATP (POTENTIAL).  
 FT METAL 284 298 MANGANESE 1 (BY SIMILARITY).  
 FT METAL 298 298 MANGANESE 1 AND 2 (BY SIMILARITY).  
 FT METAL 300 300 MANGANESE 2 (BY SIMILARITY).  
 FT METAL 820 820 MANGANESE 3 (BY SIMILARITY).  
 FT METAL 832 832 MANGANESE 3 (BY SIMILARITY).  
 SQ SEQUENCE 1058 AA; 117451 MW; ED7037AF77C1E39F CRC64;  
 Query Match 73.1%; Score 38; DB 1; Length 1058;  
 Best Local Similarity 60.0%; Pred. No. 6.2;  
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
 QY 2 EVVPGXGMSYS 11  
 Db 190 EIVPGLNYS 199  
 I::I I::I I::I  
 RESULT 2  
 SULH\_SCHPO STANDARD; PRT; 877 AA.  
 ID 074377;  
 AC 074377;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Probable sulfate permease C3H7.02.  
 GN SPAC3H7.02.  
 OS Schizosaccharomyces pombe (Fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OC Schizosaccharomycetes.  
 OX NCBI\_TaxID=4896;  
 RN [1]  
 RP SEQUENCE FROM N.A.

STRAIN-972;  
 MEDLINE=21848401; PubMed=11859360;  
 RC WOOD V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,  
 RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,  
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,  
 RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,  
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,  
 RA Holroyd S., Howarth T., Howarth S., Huckle E.J., Hunt S., Jagels K.,  
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,  
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,  
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,  
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,  
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,  
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
 RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,  
 RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,  
 RA Gabel C., Fuchs M., Fritzc C., Holzer E., Moestl D., Hilbert H.,  
 RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,  
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,  
 RA Goifeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,  
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
 RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,  
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
 RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,  
 RA Spakovski G.V., Ussery D., Barrell B.G., Nurse P.;  
 RT "The genome sequence of Schizosaccharomyces pombe".  
 RL Nature 415:871-880(2002).  
 CC -1- FUNCTION: HIGH AFFINITY UPTAKE OF SULFATE INTO THE CELL (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
 CC -1- SIMILARITY: BELONGS TO THE SLC26A FAMILY OF TRANSPORTERS.  
 CC -1- SIMILARITY: CONTAINS 1 STAS DOMAIN.  
 -----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 EMBL; AL031261; CAA20298.1;  
 InterPro; IPR002645; STAS.  
 InterPro; IPR001902; Sulfate\_transp.  
 Pfam; PF00916; Sulfate\_transp; 1.  
 Pfam; PF01740; STAS; 1.  
 TIGRFAMS; TIGR00815; sulp; 1.  
 PROSITE; PS01130; SLC26A; 1.  
 PROSITE; PS08001; STAS; 1.  
 Transport; Transmembrane.  
 FT TRANSMEM 133 153 POTENTIAL.  
 FT TRANSMEM 161 181 POTENTIAL.  
 FT TRANSMEM 186 206 POTENTIAL.  
 FT TRANSMEM 221 241 POTENTIAL.  
 FT TRANSMEM 243 263 POTENTIAL.  
 FT TRANSMEM 292 312 POTENTIAL.  
 FT TRANSMEM 329 349 POTENTIAL.  
 FT TRANSMEM 384 404 POTENTIAL.  
 FT TRANSMEM 424 444 POTENTIAL.  
 FT TRANSMEM 461 481 POTENTIAL.  
 FT TRANSMEM 484 504 POTENTIAL.  
 FT TRANSMEM 518 538 POTENTIAL.  
 FT TRANSMEM 543 563 POTENTIAL.  
 FT DOMAIN 594 747 STAS.  
 SQ SEQUENCE 877 AA; 96373 MW; 5695A8493371E43 CRC64;  
 Query Match 71.2%; Score 37; DB 1; Length 877;  
 Best Local Similarity 77.8%; Pred. No. 8.3;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 3 VVPGXGMSYS 11  
 I::I I::I I::I



Db 148 VVPQGMSSYA 156

RESULT 3

Y11K\_TYDVA STANDARD; PRT; 102 AA.

AC P31619;

DT 01-JUL-1993 (Rel. 26, Created)

DT 01-JUL-1993 (Rel. 26, Last sequence update)

DT 01-OCT-1993 (Rel. 27, Last annotation update)

DE Hypothetical 11.2 kDa protein.

GN V1.

OS Tobacco yellow dwarf virus (strain Australia) (TYDV).

OC Viruses; ssDNA viruses; Geminiviridae; Mastrevirus.

OX NCBI\_TaxID=31599;

RN [1]

RP MEDLINE=92188538; PubMed=1546458;

RA Morris B.A.M., Richardson K.A., Haley A., Zhan X., Thomas J.E.;

RT "The nucleotide sequence of the infectious cloned DNA component of tobacco yellow dwarf virus reveals features of geminiviruses infecting monocotyledonous plants.";

RL Virology 187:633-642(1992).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; M81103; AAA47947.1; -

DR PIR; A42452; A42452.

DR InterPro; IPR002621; Gemini\_mov.

DR Pfam; PF01708; Gemini\_mov; 1.

KW Hypothetical protein.

KW Hypothetical protein.

SQ SEQUENCE 102 AA; 11178 MW; A40ECF1E0AF55B67 CRC64;

Query Match 69.2%; Score 36; DB 1; Length 102;

Best Local Similarity 60.0%; Pred. No. 1.4;

Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Oy 2 EVVPXGMSYS 11

Db 7 QVFPSSGINYS 16

RESULT 4

Y1A9\_CLOAB STANDARD; PRT; 1498 AA.

AC Q04351;

DT 01-FEB-1994 (Rel. 28, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Hypothetical protein CAC3709.

GN CAC3709.

OS Clostridium acetobutylicum.

OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae; Clostridium.

OX NCBI\_TaxID=1488;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-ATCC 824 / DSM 792 / VKM B-1787;

RX MEDLINE=21359325; PubMed=11466286;

RA Neelling J., Breton G., Omelchenko M.V., Makarova K.S., Zeng Q., Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I., Tatusov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J., Bennett G.N., Koonin E.V., Smith D.R.;

RA "Genome sequence and comparative analysis of the solvent-producing bacterium Clostridium acetobutylicum.";

RT J. Bacteriol. 183:4823-4838(2001).

RN [2]

RP SEQUENCE OF 1-108 FROM N.A.

RC STRAIN-ATCC 824 / DSM 792 / VKM B-1787;

RX MEDLINE=93273706; PubMed=8501044;

RA Sauer U., Duerre P.;

RT "Sequence and molecular characterization of a DNA region encoding a small heat shock protein of Clostridium acetobutylicum.";

RL J. Bacteriol. 175:3394-3400(1993).

CC -1- SIMILARITY: BELONGS TO THE FTSK/SPOIIE FAMILY.

CC -1- CAUTION: Ref. 2 sequence differs from that shown due to frameshifts in positions 76 and 106.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; AE007866; AAK81629.1; -

DR EMBL; X65276; CAA46379.1; ALT\_FRAME.

DR InterPro; IPR002543; FtsK\_SpoiIE.

DR Pfam; PF01580; FtsK\_SpoiIE; 2.

KW Hypothetical protein; ATP-binding; Complete proteome.

FT NP\_BIND 675 682 ATP (POTENTIAL).

SQ SEQUENCE 1498 AA; 168968 MW; FF42037A335A9649 CRC64;

Query Match 69.2%; Score 36; DB 1; Length 1498;

Best Local Similarity 60.0%; Pred. No. 23;

Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Oy 1 EEVVPXGMSY 10

Db 1276 EQKIPMGMSY 1285

RESULT 5

ZEPI1\_HUMAN STANDARD; PRT; 2717 AA.

ID ZEPI1\_HUMAN

AC P15822;

DT 01-APR-1990 (Rel. 14, Created)

DT 01-APR-1990 (Rel. 14, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Zinc finger protein 40 (Human immunodeficiency virus type I enhancer-binding protein 1) (HIV-Ep1) (Major histocompatibility complex binding protein 1) (MBP-1) (Positive regulatory domain II binding factor 1) (PRDII-BF1).

DE (PRDII-BF1).

GN HIVEP1 OR ZNF40.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=90169514; PubMed=2106471;

RA Fan C.M., Maniatis T.;

RT "A DNA-binding protein containing two widely separated zinc finger motifs that recognize the same DNA sequence.";

RL Genes Dev. 4:29-42(1990).

RN [2]

RP STRUCTURE BY NMR OF 2113-2142.

RX MEDLINE=91064333; PubMed=2248949;

RA Omichinski J.G., Clore G.M., Appella E., Sakaguchi K., Gronenborn A.M.;

RT "High-resolution three-dimensional structure of a single zinc finger from a human enhancer binding protein in solution.";

RL Biochemistry 29:9324-9334(1990).

RN [3]

RP STRUCTURE BY NMR OF 2087-2142.

RX MEDLINE=92232684; PubMed=1567844;

RA Omichinski J.G., Clore G.M., Robien M., Sakaguchi K., Appella E., Gronenborn A.M.;

RT "High-resolution solution structure of the double Cys2His2 zinc

RT finger from the human enhancer binding protein MBP-1.1";  
 CC Biochemistry 31:3907-3917(1992).  
 CC -1- FUNCTION: THIS PROTEIN SPECIFICALLY BINDS TO THE DNA SEQUENCE  
 CC 5'-GGGACTTCC-3' WHICH IS FOUND IN THE ENHANCER ELEMENTS OF  
 CC NUMEROUS VIRAL PROMOTERS SUCH AS THOSE OF SV40, CMV, OR HIV1.  
 CC IN ADDITION, RELATED SEQUENCES ARE FOUND IN THE ENHANCER ELEMENTS  
 CC OF A NUMBER OF CELLULAR PROMOTERS, INCLUDING THOSE OF THE CLASS I  
 CC MHC, INTERLEUKIN-2 RECEPTOR, AND INTERFERON-BETA GENES. IT MAY ACT  
 CC IN 1-CELL ACTIVATION.  
 CC -1- SUBCELLULAR LOCATION: Nuclear.  
 CC -1- INDUCTION: BY MITOGEN AND PHORBOL ESTER.  
 CC -1- DOMAIN: CONTAINS TWO SETS OF 2 ZINC-FINGERS, WHICH ARE WIDELY  
 CC SEPARATED AND RECOGNIZE THE SAME DNA SEQUENCE. THERE IS A FIFTH  
 CC ZINC-FINGER IN-BETWEEN.  
 CC -1- SIMILARITY: STRONG, TO HIVP2.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X51435; CAA35798.1; -;  
 DR PIR; A34203; A34203.  
 DR PDB; 3ZNF; 15-JAN-92.  
 DR PDB; 4ZNF; 15-JAN-92.  
 DR PDB; 1BRO; 31-OCT-93.  
 DR TRANSFAC; T00497; -;  
 DR Genew; HGNC:4920; HIVEP1.  
 DR MIM; 194540; -;  
 DR InterPro; IPR000822; Znf\_C2H2.  
 DR Pfam; PF00096; zf-C2H2; 5.  
 DR PRINTS; PR00048; ZINC\_FINGER.  
 DR SMART; SM00355; Znf\_C2H2; 4.  
 DR PROSITE; PS00028; ZINC\_FINGER\_C2H2\_1; 4.  
 DR PROSITE; PS00157; ZINC\_FINGER\_C2H2\_2; 4.  
 KW Transcription regulation; Zinc-finger; Metal-binding; DNA-binding;  
 KW Nuclear protein; Repeat; 3D-structure.  
 FT DOMAIN 406 456  
 FT ZN\_FING 406 428  
 FT ZN\_FING 434 456  
 FT DOMAIN 803 806  
 FT ZN\_FING 808 981  
 FT ZN\_FING 958 981  
 FT DOMAIN 2087 2139  
 FT ZN\_FING 2087 2109  
 FT ZN\_FING 2115 2139  
 FT STRAND 2088 2088  
 FT TURN 2090 2092  
 FT STRAND 2095 2095  
 FT HELIX 2099 2108  
 FT TURN 2109 2109  
 FT STRAND 2115 2116  
 FT STRAND 2123 2124  
 FT HELIX 2127 2135  
 SQ SEQUENCE 2717 AA; 297217 MW; D45D3CA951FEA561 CRC64;  
 Query Match 67.3%; Score 35; DB 1; Length 2717;  
 Best Local Similarity 66.7%; Pred. NO. 70;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 QY 3 VVPXGMSYS 11  
 ||| |::|  
 Db 2405 VVPAGLTYS 2413  
 RESULT 6  
 CY14\_NEUCR  
 ID CY14\_NEUCR STANDARD; PRT; 788 AA.  
 AC P23622;  
 DT 01-NOV-1991 (Rel. 20, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Sulfate permease II.  
 GN CYS-14.  
 OS Neurospora crassa.  
 CC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
 CC Sordariales; Sordariaceae; Neurospora.  
 CC NCBI\_TaxID=5141;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91129256; PubMed=18251178;  
 RA Ketter J.S., Jarai G., Fu Y.-H., Marzluf G.A.;  
 RT "Nucleotide sequence, messenger RNA stability, and DNA recognition  
 RT elements of cys-14, the structural gene for sulfate permease II in  
 RT Neurospora crassa.";  
 RL Biochemistry 30:1780-1787(1991).  
 RN [2]  
 RP PROBABLE REVISIONS.  
 RX MEDLINE=94188926; PubMed=8140616;  
 RA Sandal N.N., Marcker K.A.;  
 RT "Similarities between a soybean nodulin, Neurospora crassa sulphate  
 RT permease II and a putative human tumour suppressor.";  
 RL Trends Biochem. Sci. 19:19-19(1994).  
 CC -1- FUNCTION: UPTAKE OF SULFATE INTO THE CELL.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -1- INDUCTION: Highly expressed, but only in cells subject to sulfur  
 CC limitation, and it is turned on by the positive-acting Cys-3  
 CC sulfur regulatory protein.  
 CC -1- MISCELLANEOUS: SULFATE PERMEASE II IS MAINLY FOUND IN MYCELIA.  
 CC -1- SIMILARITY: BELONGS TO THE SLC26A FAMILY OF TRANSPORTERS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M59167; AAA33615.1; ALT\_SEQ.  
 DR PIR; A37956; A37956.  
 DR InterPro; IPR001902; Sulfate\_transp.  
 DR Pfam; PF00916; Sulfate\_transp; 1.  
 DR TIGRFAMs; TIGR00815; sulp; 1.  
 DR PROSITE; PS01130; SLC26A; 1.  
 KW Transport; Transmembrane; Glycoprotein.  
 FT TRANSMEM 71 91  
 FT TRANSMEM 103 123  
 FT TRANSMEM 128 148  
 FT TRANSMEM 171 191  
 FT TRANSMEM 193 213  
 FT TRANSMEM 271 291  
 FT TRANSMEM 326 346  
 FT TRANSMEM 363 383  
 FT TRANSMEM 451 471  
 FT TRANSMEM 474 494  
 FT CARBOHYD 23 23  
 FT CARBOHYD 578 578  
 SQ SEQUENCE 788 AA; 87864 MW; 4FC604B60798CE77 CRC64;  
 Query Match 65.4%; Score 34; DB 1; Length 788;  
 Best Local Similarity 66.7%; Pred. NO. 32;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 QY 3 VVPXGMSYS 11  
 ||| |::|  
 Db 90 VVPQGMAYA 98  
 RESULT 7  
 A10C\_HUMAN  
 ID A10C\_HUMAN STANDARD; PRT; 1499 AA.  
 AC O60312; Q96914;  
 DT 30-MAY-2000 (Rel. 39, Created)

15-JUN-2002 (Rel. 41, Last sequence update)  
 15-JUN-2002 (Rel. 41, Last annotation update)  
 Potential phospholipid-transporting ATPase VC (EC 3.6.3.1) (ATPVC)  
 (Aninophospholipid translocase VC).  
 ATP10C OR ATPVC OR KIAA0566.  
 OS Homo sapiens (Human).  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 NCBI\_TaxID=9606;  
 [1]  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE=21225279; PubMed=11326269;  
 RA Meguro M., Kashiwagi A., Mitsuya K., Nakao M., Kondo I., Saitoh S.,  
 Oshimura M.;  
 RA "A novel maternally expressed gene, ATP10C, encodes a putative  
 aminophospholipid translocase associated with Angelman syndrome.";  
 Nat. Genet. 28:19-20(2001).  
 [2]  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE=21313119; PubMed=11353404;  
 RA Herzig L.B.K., Kim S.-J., Cook E.H. Jr., Ledbetter D.H.;  
 RA "The human aminophospholipid-transporting ATPase gene ATP10C maps  
 adjacent to UBE3A and exhibits similar imprinted expression.";  
 Am. J. Hum. Genet. 68:1501-1505(2001).  
 [3]  
 RN SEQUENCE OF 337-1499 FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=98290545; PubMed=9628581;  
 RA Nagase T., Ishikawa K.-I., Miyajima N., Tanaka A., Kotani H.,  
 Nomura N., Ohara O.;  
 RA "Prediction of the coding sequences of unidentified human genes. IX.  
 The complete sequences of 100 new cDNA clones from brain which can  
 code for large proteins in vitro.";  
 DNA Res. 5:31-39(1998).  
 CC -!- CATALYTIC ACTIVITY: ATP + H(2)O -> ADP + phosphate.  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).  
 CC -!- TISSUE SPECIFICITY: Wide expression, with highest levels in  
 kidney, followed by lung, brain, prostate, testis, ovary, and  
 small intestine.  
 CC -!- DISEASE: Defects in ATP10C may be a cause of Angelman syndrome  
 (AS), also known as 'happy puppet syndrome'.  
 CC -!- SIMILARITY: BELONGS TO THE CATION TRANSPORT ATPASES FAMILY (E1-E2  
 ATPASES). SUBFAMILY IV.  
 -----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 DR EMBL; AB051358; BAB47392.1; -;  
 DR EMBL; AY029504; AAK33100.1; -;  
 DR EMBL; AY029487; AAK33100.1; JOINED.  
 DR EMBL; AY029488; AAK33100.1; JOINED.  
 DR EMBL; AY029489; AAK33100.1; JOINED.  
 DR EMBL; AY029490; AAK33100.1; JOINED.  
 DR EMBL; AY029491; AAK33100.1; JOINED.  
 DR EMBL; AY029492; AAK33100.1; JOINED.  
 DR EMBL; AY029493; AAK33100.1; JOINED.  
 DR EMBL; AY029494; AAK33100.1; JOINED.  
 DR EMBL; AY029495; AAK33100.1; JOINED.  
 DR EMBL; AY029496; AAK33100.1; JOINED.  
 DR EMBL; AY029497; AAK33100.1; JOINED.  
 DR EMBL; AY029498; AAK33100.1; JOINED.  
 DR EMBL; AY029499; AAK33100.1; JOINED.  
 DR EMBL; AY029500; AAK33100.1; JOINED.  
 DR EMBL; AY029501; AAK33100.1; JOINED.  
 DR EMBL; AY029502; AAK33100.1; JOINED.  
 DR EMBL; AY029503; AAK33100.1; JOINED.  
 DR EMBL; AB011138; BAA25492.1; -;  
 DR Genew; HGNC:13547; ATP10C.

DR MIM; 605855; -;  
 DR MIM; 105830; -;  
 DR InterPro; IPR001757; ATPase\_E1-E2  
 DR InterPro; IPR001454; Hlgase/hydrilase.  
 DR Pfam; PF00702; Hydrolase; 1.  
 DR PRINTS; PR00119; CATAPASE.  
 DR PROSITE; PS00154; ATPase\_E1\_E2; 1.  
 DR Hydrolase; Transmembrane; Phosphorylation; Magnesium; ATP-binding;  
 KW Multigene family.  
 FT DOMAIN 1 86 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 87 106 POTENTIAL.  
 FT DOMAIN 107 110 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 111 128 POTENTIAL.  
 FT DOMAIN 129 309 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 310 332 POTENTIAL.  
 FT DOMAIN 337 362 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 363 384 POTENTIAL.  
 FT DOMAIN 385 1087 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 1088 1108 POTENTIAL.  
 FT DOMAIN 1109 1119 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 1120 1140 POTENTIAL.  
 FT DOMAIN 1141 1170 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 1171 1192 POTENTIAL.  
 FT DOMAIN 1193 1199 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 1200 1222 POTENTIAL.  
 FT DOMAIN 1223 1228 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 1229 1249 POTENTIAL.  
 FT DOMAIN 1250 1267 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 1268 1292 POTENTIAL.  
 FT DOMAIN 1293 1499 CYTOPLASMIC (POTENTIAL).  
 FT MOD\_RES 427 427 PHOSPHORYLATION (BY SIMILARITY).  
 FT METAL 1031 1031 MAGNESIUM (BY SIMILARITY).  
 FT METAL 1035 1035 MAGNESIUM (BY SIMILARITY).  
 FT DOMAIN 467 470 POLY-GLU.  
 FT CONFLICT 388 388 Q -> R (IN REF. 3).  
 SQ SEQUENCE 1499 AA; 167687 MW; D4996A4D0635A68D CRC64;  
 Query Match 65.4%; Score 34; DB 1; Length 1499;  
 Best Local Similarity 72.7%; Pred. No. 62;  
 Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 BEVPXGMSYS 11  
 Db 469 BEVPXGMSYS 479  
 -----  
 RESULT 8  
 ID RL20\_MYCPU STANDARD; PRT; 116 AA.  
 AC Q98QV0;  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE 50S ribosomal protein L20.  
 GN RPLT OR MYPU\_2610.  
 OS Mycoplasma pulmonis.  
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.  
 OX NCBI\_TaxID=2107;  
 RN [1]  
 RC SEQUENCE FROM N.A.  
 RP STRAIN=UAB CTIP;  
 RX MEDLINE=21267165; PubMed=11353084;  
 RA Chambaud I., Heilig R., Ferris S., Barbe V., Samson D., Galisson F.,  
 Moszer I., Dybvig K., Wroblewski H., Viari A., Rocha E.P.C.,  
 Blanchard A.;  
 RA "The complete genome sequence of the murine respiratory pathogen  
 Mycoplasma pulmonis.";  
 Nucleic Acids Res. 29:2145-2153(2001).  
 CC -!- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 23S RIBOSOMAL RNA AND IS  
 NECESSARY TO THE IN VITRO ASSEMBLY PROCESS OF THE 50S RIBOSOMAL  
 SUBUNIT; IT IS NOT INVOLVED IN THE PROTEIN SYNTHESIZING FUNCTIONS  
 OF THAT SUBUNIT (BY SIMILARITY).  
 CC -!- SIMILARITY: BELONGS TO THE L20P FAMILY OF RIBOSOMAL PROTEINS.

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AL445563; CAC13434.1; -  
 CC Mypulist; MYP0\_2610; -  
 CC InterPro; IPR001081; Ribosomal\_L20.  
 CC Pfam; PF00453; Ribosomal\_L20; 1.  
 CC PRINTS; PR00062; RIBOSOMAL\_L20.  
 CC ProDom; PD002389; Ribosomal\_L20; 1.  
 CC TIGRFAMs; TIGR01032; rplT\_bact; 1.  
 CC PROSITE; PS00937; RIBOSOMAL\_L20; 1.  
 CC Ribosomal protein; rRNA-binding; Complete proteome.  
 CC SQ SEQUENCE 116 AA; 13565 MW; C59C748901B18F14 CRC64;  
 CC -----  
 CC Query Match 63.5%; Score 33; DB 1; Length 116;  
 CC Best Local Similarity 77.8%; Pred. No. 7;  
 CC Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 CC -----  
 CC QY 3 VVPXGMSYS 11  
 CC | | | | |  
 CC Db 68 VRPLGMSYS 76  
 CC -----  
 CC RESULT 9  
 CC Y990\_CAMJE  
 CC ID Y990\_CAMJE STANDARD; PRT; 253 AA.  
 CC AC P45489; Q9PNV0;  
 CC DT 01-NOV-1995 (Rel. 32, Created)  
 CC DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 CC DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 CC DE Hypothetical protein Cj0990C.  
 CC GN Cj0990C.  
 CC OS Campylobacter jejuni.  
 CC OC Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group;  
 CC CC Campylobacter.  
 CC NCBI\_TaxID=197;  
 CC [1]  
 CC SEQUENCE FROM N.A.  
 CC STRAIN=NCTC 11168;  
 CC MEDLINE=20150912; PubMed=10688204;  
 CC RA Parkhill J., Wren B.W., Mungall K., Kettle J.M., Churcher C.,  
 CC Basham D., Chillingworth T., Davies R.M., Feltwell T., Holtroyd S.,  
 CC Jagers K., Karlyshev A.V., Moule S., Pallen M.J., Penn C.W.,  
 CC Quail M.A., Rajandream M.A., Rutherford K.M., van Vliet A.H.M.,  
 CC Whitehead S., Barrell B.G.;  
 CC "The genome sequence of the food-borne pathogen Campylobacter jejuni  
 CC reveals hypervariable sequences.";  
 CC Nature 403:665-668(2000).  
 CC [2]  
 CC SEQUENCE OF 160-253 FROM N.A.  
 CC STRAIN=ATCC 43431 / TGH 9011;  
 CC MEDLINE=95247673; PubMed=7730270;  
 CC Hani E.K., Chan V.L.;  
 CC "Expression and characterization of Campylobacter jejuni  
 CC benzoylglycine amidohydrolase (Hippuricase) gene in Escherichia  
 CC coli.";  
 CC J. Bacteriol. 177:2396-2402(1995).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AL139076; CAB73246.1; -

DR EMBL; Z36940; CAA85392.1; -  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 253 AA; 29783 MW; F96D3FF3265F8A6A CRC64;  
 CC -----  
 CC Query Match 63.5%; Score 33; DB 1; Length 253;  
 CC Best Local Similarity 55.6%; Pred. No. 16;  
 CC Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
 CC -----  
 CC QY 2 EVVPXGMSY 10  
 CC | | | | |  
 CC Db 185 DIFFSGMSY 193  
 CC -----  
 CC RESULT 10  
 CC AROA\_VIBCH  
 CC ID AROA\_VIBCH STANDARD; PRT; 426 AA.  
 CC AC Q9RBO0;  
 CC DT 15-JUN-2002 (Rel. 41, Created)  
 CC DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 CC DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 CC DE 3-phosphorylshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) (5-  
 CC enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) (EPSPS).  
 CC GN AROA OR VC1732.  
 CC OS Vibrio cholerae.  
 CC OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.  
 CC NCBI\_TaxID=666;  
 CC [1]  
 CC SEQUENCE FROM N.A.  
 CC STRAIN=El Tor N16961 / Serotype O1;  
 CC MEDLINE=20406833; PubMed=10952301;  
 CC RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,  
 CC Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,  
 CC Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,  
 CC Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,  
 CC McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,  
 CC Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,  
 CC Fraser C.M.;  
 CC "DNA sequence of both chromosomes of the cholera pathogen Vibrio  
 CC cholerae.";  
 CC Nature 406:477-483(2000).  
 CC CC CATALYTIC ACTIVITY: Phosphoenolpyruvate + 3-phosphoshikimate =  
 CC phosphate + 5-O-(1-carboxyvinyl)-3-phosphoshikimate.  
 CC CC -/- PATHWAY: Aromatic amino acids biosynthesis; shikimate pathway;  
 CC sixth step.  
 CC CC -/- SUBUNIT: Monomer (By similarity).  
 CC CC -/- SUBCELLULAR LOCATION: Cytoplasmic (Probable).  
 CC CC -/- SIMILARITY: BELONGS TO THE EPSP SYNTHASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AE004251; AAF94882.1; -  
 CC TIGR; VC1732;  
 CC InterPro; IPR001986; EPSP\_synthase.  
 CC Pfam; PF00275; EPSP\_synthase; 1.  
 CC ProDom; PD001867; EPSP\_synthase; 1.  
 CC PROSITE; PS00104; EPSP\_SYNTHASE\_1; 1.  
 CC PROSITE; PS00885; EPSP\_SYNTHASE\_2; 1.  
 CC Aromatic amino acid biosynthesis; Transferase; Complete proteome.  
 CC SQ SEQUENCE 426 AA; 46101 MW; 38852D6483BFE1C3 CRC64;  
 CC -----  
 CC Query Match 63.5%; Score 33; DB 1; Length 426;  
 CC Best Local Similarity 60.0%; Pred. No. 27;  
 CC Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
 CC -----  
 CC QY 1 EEVVPXGMSY 10  
 CC | | | | |  
 CC Db 223 EFVIPAGQSY 232  
 CC -----

Query Match 63.5%; Score 33; DB 1; Length 466;  
 Best Local Similarity 50.0%; Pred. No. 30;  
 Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10  
 : : : : :  
 Db 98 DSAIPGMSY 107

## RESULT 12

ISR2\_HUMAN STANDARD; PRT; 478 AA.  
 ID GSR2\_HUMAN STANDARD; PRT; 478 AA.

AC Q9NZM5; Q9NPP1; Q9NPR4; Q9UFI2; Q9BTC6; Q9HAX6;  
 DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Glioma tumor suppressor candidate region gene 2 protein (p60).

GN GLTSCR2.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.

RC MEDLINE=20175430; PubMed=10708517;

RA Smith J.S., Tachibana I., Pohl U., Lee H.K., Thanarajasingam U.,

RA Portier B.P., Ueki K., Billings S., Ramaswamy S., Mohrenweiser H.W.,

RA Scheithauer B.W., Louis D.N., Jenkins R.B.,

RT "A transcript map of the chromosome 19q-Arm glioma tumor suppressor

RT region.";

RL Genomics 64:44-50(2000).

RN [2]

RP SEQUENCE FROM N.A.

RC TISSUE=Muscle;

RA Strausberg R.;

RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE OF 9-478 FROM N.A., AND SUBCELLULAR LOCATION.

RC MEDLINE=95214318; PubMed=10196275;

RA Bruni R., Fineschi B., Ogle W.O., Roizman B.;

RT "A novel cellular protein, p60, interacting with both herpes simplex

RT virus 1 regulatory proteins ICP22 and ICP0 is modified in a

RT cell-type-specific manner and is recruited to the nucleus after

RT infection.";

RL J. Virol. 73:3810-3817(1999).

RN [4]

RP SEQUENCE OF 12-478 FROM N.A.

RC Andreu N., Estivill X., Escarceller M., Sumoy L.;

RA Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.

RN [5]

RP SEQUENCE OF 218-477 FROM N.A.

RC TISSUE=Testis;

RA Poustka A., Klein M., Mewes H.-W., Gassenhuber J., Wiemann S.;

RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.

CC -!- SUBUNIT: Interacts with HSP-1 early proteins ICP22 and ICP0.

CC -!- SUBCELLULAR LOCATION: Nuclear.

CC -!- TISSUE SPECIFICITY: Expressed at high levels in heart and

CC pancreas, moderate levels in placenta, liver, skeletal muscle, and

CC kidney, and low levels in brain and lung.

CC -!- SIMILARITY: BELONGS TO THE GLTSCR2 FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

CC EMBL; AF182076; AAF62873.1; -.

DR EMBL; BC004229; AAH04229.1; -.

DR EMBL; BC006311; AAH06311.1; -.

DR

## RESULT 11

CC37\_SCHPO STANDARD; PRT; 466 AA.

ID CC37\_SCHPO STANDARD; PRT; 466 AA.

AC O94740;

DT 15-JUN-2002 (Rel. 41, Created)

DT 15-JUN-2002 (Rel. 41, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Hsp90 co-chaperone Cdc37 (Hsp90 chaperone protein kinase-targeting

DE subunit) (Cell division control protein 37).

GN CDC37 OR SPAC9B6.10.

OS Schizosaccharomyces pombe (Fission yeast).

OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;

OC Schizosaccharomycetales; Schizosaccharomycetaceae;

OC Schizosaccharomycetes.

OX NCBI\_TaxID=4896;

RN [1]

RP SEQUENCE FROM N.A.

RC Westwood P.K., Preston N.C., Pantos P.A.;

RT "Schizosaccharomycetes pombe cdc37 gene.";

RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=972;

RA MEDLINE=21848401; PubMed=11859360;

RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,

RA Sgouras J., Feat N., Hayes J., Baker S., Basham D., Bowman S.,

RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,

RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,

RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,

RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,

RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,

RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odeh J.,

RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,

RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,

RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,

RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,

RA Welljens I., Vanslaers E., Rieger M., Schaefer M., Mueller-Auer S.,

RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,

RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,

RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,

RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,

RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,

RA Lucas M., Rochet M., Gallard J., Tardieu V.A., Garzon A., Thode G.,

RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,

RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,

RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,

RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;

RT "The genome sequence of Schizosaccharomyces pombe.";

RL Nature 415:871-880(2002).

CC -!- FUNCTION: With Hsp90 it forms a complex that binds to several

CC kinases, resulting in stabilization and promotion of their

CC activity (By similarity).

CC -!- SUBUNIT: Forms a complex with Hsp90. Interacts with a number of

CC kinases (By similarity).

CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).

CC -!- SIMILARITY: BELONGS TO THE CDC37 FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

CC EMBL; AJ132377; CAB38758.1; -.

DR EMBL; AJ132376; CAB38757.1; -.

DR EMBL; AL049769; CAB42371.2; -.

KW Chapterone; Cell division; Cell cycle.

SEQUENCE 466 AA; 52554 MW; 647238B34CABB3C5 CRC64;

DR

DR EMBL; BC010095; AAH10095.1; -  
 DR EMBL; AF296124; AG30413.1; -  
 DR EMBL; AL359335; CAB94786.1; -  
 DR EMBL; AL359336; CAB94787.1; -  
 DR EMBL; AL122063; CAB59242.1; -  
 DR SWISS-2DPAGE; Q9NZM5; HUMAN.  
 DR Genew; HGNC:4333; GLTSCR2.  
 DR MIM; 605691; -  
 KW Nuclear protein; Polymorphism.  
 FT VARIANT 389 389 R -> Q.  
 FT /FTID=VAR\_011486.  
 FT GGS -> HEG (IN REF. 2; AAH04229).  
 FT G -> R (IN REF. 3).  
 FT RKEQLWEKLAGSGELPREVRRQAARLLNPSATRAKPGQD  
 FT TVERP -> SGRSSYGRSWPSSRASSPGGAQGPSVAQPFNC  
 FT KGNPAPGHRIAA (IN REF. 3).  
 FT SDNPLDRPLVGODEFFLE -> LNNPDKPVVWPGCLFPF  
 FT (IN REF. 3).  
 FT A -> S (IN REF. 2; AAH04229).  
 FT D -> H (IN REF. 3).  
 FT PEGNILDREKFSQRRNMIEPRERAKFKRKYKVLVEKRAF  
 FT REIQ -> VLTVSCRGAPCPVMTPLSLPPVPRGYGRHHGCP  
 FT WAGPVGMPRG (IN REF. 5).  
 FT EGNILDRKFSQRRNMIEPRERAKFKRKYKVLVEKRAF  
 FT ETQL -> RQHSFETGSAFRGGI (IN REF. 3).  
 FT SEQUENCE 478 AA; 54417 MW; 7F18923E348CB52B CRC64;  
 Query Match 63.5%; Score 33; DB 1; Length 478;  
 Best Local Similarity 60.0%; Pred. No. 30;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
 QY 2 EVVPGMSYS 11  
 |||||  
 DB 239 EVAPAGASIN 248  
 RESULT 13  
 BCN5\_CLOPE  
 ID BCN5\_CLOPE STANDARD; PRT; 890 AA.  
 AC P08696;  
 DT 01-JAN-1988 (Rel. 06, Created)  
 DT 01-JAN-1988 (Rel. 06, Last sequence update)  
 DT 01-OCT-1994 (Rel. 30, Last annotation update)  
 DE Bacteriocin BCN5.  
 GN BCN.  
 OS Clostridium perfringens.  
 OG Plasmid pIP404.  
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;  
 OC Clostridium.  
 OX NCBI\_TaxID=1502;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CPN50;  
 RX MEDLINE=88336297; PubMed=2901768;  
 RA Garnier T., Cole S.T.;  
 RT "Complete nucleotide sequence and genetic organization of the  
 RT bacteriocinogenic plasmid, pIP404, from Clostridium perfringens.";  
 RL Plasmid 19:134-150(1988).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CPN50;  
 RX MEDLINE=87057020; PubMed=2877971;  
 RA Garnier T., Cole S.T.;  
 RT "Characterization of a bacteriocinogenic plasmid from Clostridium  
 RT perfringens and molecular genetic analysis of the  
 RT bacteriocin-encoding gene.";  
 RL J. Bacteriol. 168:1189-1196(1986).  
 RN [3]  
 RP SEQUENCE OF 1-14 FROM N.A.  
 RC STRAIN=CPN50;  
 RX MEDLINE=89039249; PubMed=2460717;  
 RA Garnier T., Cole S.T.;  
 RT "Studies of UV-inducible promoters from Clostridium perfringens in

vivo and in vitro.";  
 RL Mol. Microbiol. 2:607-614(1988).  
 CC -!- FUNCTION: MAY FUNCTION AS AN IONOPHORE.  
 CC -!- INDUCTION: BY UV IRRADIATION.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M14481; AAA98248.1; -  
 DR EMBL; M32882; AAA98249.1; -  
 DR PIR; A30481; A30481.  
 DR InterPro; IPR003646; SH3\_bac.  
 DR SMART; SM00287; SH3b; 3.  
 DR Antibiotic; Bacteriocin; Plasmid.  
 KW DOMAIN 815 869 HYDROPHOBIC.  
 FT SEQUENCE 890 AA; 96699 MW; F4E5E8971C31C6C6 CRC64;  
 SQ  
 Query Match 63.5%; Score 33; DB 1; Length 890;  
 Best Local Similarity 66.7%; Pred. No. 58;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 QY 2 EVVPGMSY 10  
 |||||  
 DB 170 EVVPGFTY 178  
 RESULT 14  
 LTBS\_HUMAN  
 ID LTBS\_HUMAN STANDARD; PRT; 1394 AA.  
 AC P22064;  
 DT 01-AUG-1991 (Rel. 19, Created)  
 DT 01-AUG-1991 (Rel. 19, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Latent transforming growth factor beta binding protein 1s precursor  
 DE (Transforming growth factor beta-1 binding protein 1) (TGF-beta1-BP-  
 DE 1).  
 GN LTBP1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RC TISSUE=Fibroblast, and Platelet;  
 RX MEDLINE=90275601; PubMed=2350783;  
 RA Kanzaki T., Olofsson A., Moren A., Wernstedt C., Hellman U.,  
 RA Miyazono K., Claesson-Welsh L., Heldin C.-H.;  
 RT "TGF-beta 1 binding protein: a component of the large latent complex  
 RT of TGF-beta 1 with multiple repeat sequences.";  
 RL Cell 61:1051-1061(1990).  
 CC -!- SUBUNIT: THE LARGE LATENT COMPLEX OF TGF-BETA1 FROM PLATELETS IS  
 CC COMPOSED OF THE TGF-BETA1 MOLECULE NONCOVALENTLY ASSOCIATED WITH  
 CC A DISULFIDE-BONDED COMPLEX OF A DIMER OF THE N-TERMINAL PROPEPTIDE  
 CC OF THE TGF-BETA1 PRECURSOR AND A THIRD COMPONENT DENOTED TGF-  
 CC BETA1-BP. TGF-BETA1-BP DOES NOT BIND DIRECTLY TO ACTIVE TGF-BETA1.  
 CC -!- ALTERNATIVE PRODUCTS: 2 isoforms; a short form (shown here) and a  
 CC long form (AC 014766); are produced by alternative splicing.  
 CC -!- PTM: CONTAINS HYDROXYLATED ASPARAGINE RESIDUES.  
 CC -!- PTM: The N-terminus is blocked.  
 CC -!- SIMILARITY: CONTAINS 16 EGF-LIKE DOMAINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

EMBL; M34057; AAA61160.1; -  
 PIR; A35626; A35626.  
 HSSP; P00750; 1TPG.  
 GlycoSuiteDB; P22064; -  
 GenBank; HGNC:6714; LTPB1.  
 MIM; 150390; -  
 InterPro; IPR000152; Asx\_hydroxyl.  
 InterPro; IPR000561; EGF-like.  
 InterPro; IPR001881; EGF\_CA.  
 InterPro; IPR002212; Fibrin-assoc.  
 Pfam; PF00008; EGF; 15.  
 Pfam; PF00683; TB; 4.  
 SMART; SM00179; EGF\_CA; 13.  
 SMART; SM00001; EGF\_like; 4.  
 PROSITE; PS00010; ASX\_HYDROXYL; 13.  
 PROSITE; PS00022; EGF\_1; 2.  
 PROSITE; PS01186; EGF\_2; 11.  
 PROSITE; PS01187; EGF\_CA; 15.  
 Growth factor binding; Repeat; EGF-like domain; Hydroxylation; Signal;  
 Glycoprotein; Alternative splicing.  
 SIGNAL 1 20  
 CHAIN 21 1394  
 DOMAIN 300 340  
 REPEAT 348 412  
 DOMAIN 546 587  
 DOMAIN 588 629  
 DOMAIN 630 670  
 DOMAIN 671 710  
 DOMAIN 711 751  
 DOMAIN 752 792  
 DOMAIN 793 833  
 DOMAIN 834 874  
 DOMAIN 875 916  
 DOMAIN 917 958  
 DOMAIN 959 1001  
 REPEAT 1017 1084  
 DOMAIN 1097 1139  
 REPEAT 1190 1262  
 DOMAIN 1140 1180  
 DOMAIN 1294 1334  
 DOMAIN 1335 1379  
 SITE 847 849  
 DISULFID 304 315  
 DISULFID 310 324  
 DISULFID 326 339  
 DISULFID 550 562  
 DISULFID 557 571  
 DISULFID 573 586  
 DISULFID 592 604  
 DISULFID 599 613  
 DISULFID 615 628  
 DISULFID 634 645  
 DISULFID 640 654  
 DISULFID 657 669  
 DISULFID 675 686  
 DISULFID 681 695  
 DISULFID 698 709  
 DISULFID 715 726  
 DISULFID 721 735  
 DISULFID 737 750  
 DISULFID 756 767  
 DISULFID 762 776  
 DISULFID 778 791  
 DISULFID 797 808  
 DISULFID 803 817  
 DISULFID 819 832  
 DISULFID 838 850  
 DISULFID 845 859  
 DISULFID 861 873  
 DISULFID 879 891  
 DISULFID 885 900  
 EMBL; M34057; AAA61160.1; -  
 PIR; A35626; A35626.  
 HSSP; P00750; 1TPG.  
 GlycoSuiteDB; P22064; -  
 GenBank; HGNC:6714; LTPB1.  
 MIM; 150390; -  
 InterPro; IPR000152; Asx\_hydroxyl.  
 InterPro; IPR000561; EGF-like.  
 InterPro; IPR001881; EGF\_CA.  
 InterPro; IPR002212; Fibrin-assoc.  
 Pfam; PF00008; EGF; 15.  
 Pfam; PF00683; TB; 4.  
 SMART; SM00179; EGF\_CA; 13.  
 SMART; SM00001; EGF\_like; 4.  
 PROSITE; PS00010; ASX\_HYDROXYL; 13.  
 PROSITE; PS00022; EGF\_1; 2.  
 PROSITE; PS01186; EGF\_2; 11.  
 PROSITE; PS01187; EGF\_CA; 15.  
 Growth factor binding; Repeat; EGF-like domain; Hydroxylation; Signal;  
 Glycoprotein; Alternative splicing.  
 SIGNAL 1 20  
 CHAIN 21 1394  
 DOMAIN 300 340  
 REPEAT 348 412  
 DOMAIN 546 587  
 DOMAIN 588 629  
 DOMAIN 630 670  
 DOMAIN 671 710  
 DOMAIN 711 751  
 DOMAIN 752 792  
 DOMAIN 793 833  
 DOMAIN 834 874  
 DOMAIN 875 916  
 DOMAIN 917 958  
 DOMAIN 959 1001  
 REPEAT 1017 1084  
 DOMAIN 1097 1139  
 REPEAT 1190 1262  
 DOMAIN 1140 1180  
 DOMAIN 1294 1334  
 DOMAIN 1335 1379  
 SITE 847 849  
 DISULFID 304 315  
 DISULFID 310 324  
 DISULFID 326 339  
 DISULFID 550 562  
 DISULFID 557 571  
 DISULFID 573 586  
 DISULFID 592 604  
 DISULFID 599 613  
 DISULFID 615 628  
 DISULFID 634 645  
 DISULFID 640 654  
 DISULFID 657 669  
 DISULFID 675 686  
 DISULFID 681 695  
 DISULFID 698 709  
 DISULFID 715 726  
 DISULFID 721 735  
 DISULFID 737 750  
 DISULFID 756 767  
 DISULFID 762 776  
 DISULFID 778 791  
 DISULFID 797 808  
 DISULFID 803 817  
 DISULFID 819 832  
 DISULFID 838 850  
 DISULFID 845 859  
 DISULFID 861 873  
 DISULFID 879 891  
 DISULFID 885 900







```

FT DISULFID 994 1005 BY SIMILARITY.
FT DISULFID 1000 1014 BY SIMILARITY.
FT DISULFID 1017 1028 BY SIMILARITY.
FT DISULFID 1034 1045 BY SIMILARITY.
FT DISULFID 1040 1054 BY SIMILARITY.
FT DISULFID 1056 1069 BY SIMILARITY.
FT DISULFID 1075 1086 BY SIMILARITY.
FT DISULFID 1081 1095 BY SIMILARITY.
FT DISULFID 1097 1110 BY SIMILARITY.
FT DISULFID 1116 1127 BY SIMILARITY.
FT DISULFID 1132 1136 BY SIMILARITY.
FT DISULFID 1138 1151 BY SIMILARITY.
FT DISULFID 1157 1169 BY SIMILARITY.
FT DISULFID 1164 1178 BY SIMILARITY.
FT DISULFID 1180 1192 BY SIMILARITY.
FT DISULFID 1198 1210 BY SIMILARITY.
FT DISULFID 1204 1219 BY SIMILARITY.
FT DISULFID 1221 1234 BY SIMILARITY.
FT DISULFID 1240 1252 BY SIMILARITY.
FT DISULFID 1245 1261 BY SIMILARITY.
FT DISULFID 1263 1276 BY SIMILARITY.
FT DISULFID 1282 1294 BY SIMILARITY.
FT DISULFID 1289 1303 BY SIMILARITY.
FT DISULFID 1305 1319 BY SIMILARITY.
FT DISULFID 1419 1432 BY SIMILARITY.
FT DISULFID 1427 1441 BY SIMILARITY.
FT DISULFID 1443 1456 BY SIMILARITY.
FT DISULFID 1462 1473 BY SIMILARITY.
FT DISULFID 1468 1482 BY SIMILARITY.
FT DISULFID 1484 1497 BY SIMILARITY.
FT DISULFID 1616 1627 BY SIMILARITY.
FT DISULFID 1622 1636 BY SIMILARITY.
FT DISULFID 1638 1651 BY SIMILARITY.
FT DISULFID 1657 1672 BY SIMILARITY.
FT DISULFID 1667 1681 BY SIMILARITY.
FT DISULFID 1683 1696 BY SIMILARITY.
FT CARBOHYD 339 339 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 370 370 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 416 416 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 612 612 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1042 1042 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1242 1242 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1357 1357 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1712 AA; 186598 MW; 650BCCAA691FD134 CRC64;

Query Match 63.5%; Score 33; DB 1; Length 1712;
Best Local Similarity 45.5%; Pred. No. 1.le+02;
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSYS 11
:|:|:|:|:
DB 718 KEICPGMGYT 728

RESULT 18
CGD2_RAT
ID CGD2_RAT STANDARD; PRT; 288 AA.
AC Q04827;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE G1/S-specific cyclin D2 (Vin-1 proto-oncogene).
GN CCND2 OR VIN-1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93275661; PubMed=8502486;
RA Hanna Z., Jankowski M., Tremblay P., Jiang X.M., Milatovich A.,
RA Francke U., Jolicoeur P.;
RT "The Vin-1 gene, identified by provirus insertional mutagenesis, is

```

```

RT the cyclin D2.";
RL Oncogene 8:1661-1666(1993).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=95011623; PubMed=7926809;
RX Hosokawa Y., Onga T., Nakashima K.;
RT "Induction of D2 and D3 cyclin-encoding genes during promotion of the
RT G1/S transition by prolactin in rat Nb2 cells.";
RL Gene 147:249-252(1994).
CC -!- FUNCTION: ESSENTIAL FOR THE CONTROL OF THE CELL CYCLE AT THE G1/S
CC (START) TRANSITION.
CC -!- SUBUNIT: INTERACTS WITH THE CDK4 AND CDK6 PROTEIN KINASES TO FORM
CC A SERINE/THREONINE KINASE HOLOENZYME COMPLEX. THE CYCLIN SUBUNIT
CC IMPARTS SUBSTRATE SPECIFICITY TO THE COMPLEX.
CC -!- SIMILARITY: BELONGS TO THE CYCLIN FAMILY. CYCLIN D SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L09752; AAA41010.1; -
CC EMBL; D16308; BAA03815.1; -
CC InterPro; IPR004366; Cyclin.
CC InterPro; IPR004367; Cyclin_Cterm.
CC Pfam; PF00134; cyclin; 1.
CC Pfam; PF02984; cyclin_C; 1.
CC SMART; SM00385; CYCLIN; 1.
CC PROSITE; PS00292; CYCLINS; 1.
CC Cyclin; Cell cycle; Cell division; Multigene family; Proto-oncogene.
FT CONFLICT 68 68 E -> G (IN REF. 2).
FT CONFLICT 104 104 C -> V (IN REF. 2).
FT CONFLICT 232 232 T -> A (IN REF. 2).
SQ SEQUENCE 288 AA; 32826 MW; 4B522BF4E9835FC1 CRC64;

Query Match 61.5%; Score 32; DB 1; Length 288;
Best Local Similarity 60.0%; Pred. No. 29;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10
||| | | |
DB 73 EEVPEPLAMNY 82

RESULT 19
CGD2_HUMAN
ID CGD2_HUMAN STANDARD; PRT; 289 AA.
AC P30279; Q13955;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE G1/S-specific cyclin D2.
GN CCND2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92347851; PubMed=1386336;
RX Xiong Y., Manninger J., Beach D., Ward D.C.;
RT "Molecular cloning and chromosomal mapping of CCND genes encoding
RT human D-type cyclins";
RL Genomics 13:575-584(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=93205384; PubMed=8455931;
RA Palmero I., Holder A., Sinclair A.J., Dickson C., Peters G.;
RT "Cyclins D1 and D2 are differentially expressed in human B-lymphoid
RT cell lines.";

```

```

RL  Oncogene 8:1049-1054(1993).
RN  [3]
RP  SEQUENCE FROM N.A.
RA  Miyajima N.;
RL  Submitted (MAR-1993) to the EMBL/GenBank/DBJ databases.
RN  [4]
RP  SEQUENCE FROM N.A.
RA  Tissue-Bone marrow;
RL  Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
RN  [5]
RP  SEQUENCE OF 1-240 FROM N.A.
RA  Inaba T., Matsushime H., Valentine M., Roussel M.F., Sherr C.J.,
RL  "Genomic organization, chromosomal localization, and independent
RT  expression of human cyclin D genes.";
RN  Genomics 13:565-574(1992).
CC  -1- FUNCTION: ESSENTIAL FOR THE CONTROL OF THE CELL CYCLE AT THE G1/S
CC  (START) TRANSITION.
CC  -1- SUBUNIT: INTERACTS WITH THE CDK4 AND CDK6 PROTEIN KINASES TO FORM
CC  A SERINE/THREONINE KINASE HOLOENZYME COMPLEX. THE CYCLIN SUBUNIT
CC  IMPARTS SUBSTRATE SPECIFICITY TO THE COMPLEX.
CC  -1- SIMILARITY: BELONGS TO THE CYCLIN FAMILY. CYCLIN D SUBFAMILY.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; M90813; AAA51926.1; -
DR  EMBL; X68452; CAA48493.1; -
DR  EMBL; D13639; BAA02802.1; -
DR  EMBL; BC010958; AAI10958.1; -
DR  EMBL; M88083; AAA51928.1; -
DR  EMBL; M88080; AAA51928.1; JOINED.
DR  EMBL; M88081; AAA51928.1; JOINED.
DR  EMBL; M88082; AAA51928.1; JOINED.
DR  PIR; A42822; A42822.
DR  PIR; S26580; S26580.
DR  Genew; HGNC:1583; CCND2.
DR  MIM; 123833; -
DR  InterPro; IPR004366; Cyclin.
DR  InterPro; IPR004367; Cyclin_Cterm.
DR  Pfam; PF00134; cyclin; 1.
DR  Pfam; PF02984; cyclin_C; 1.
DR  SMART; SM00385; CYCLIN; 1.
DR  PROSITE; PS00292; CYCLINS; 1.
DR  Cyclin; Cell cycle; Cell division; Multigene family.
DR  CONFLICT 166 167 KL -> NV (IN REF. 5).
DR  CONFLICT 224 224 T -> H (IN REF. 5).
DR  SEQUENCE 289 AA; 33067 MW; E4E5FEF476D76D90 CRC64;

Query Match 61.5%; Score 32; DB 1; Length 289;
Best Local Similarity 60.0%; Pred. No. 29;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10
DB 74 EEVFPPLAMNY 83

RESULT 20
CGD2_MOUSE
ID CGD2_MOUSE STANDARD; PRT; 289 AA.
AC P30280;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)

```

```

DE  G1/S-specific cyclin D2.
GN  CCND2 OR CYL-2.
OS  Mus musculus (Mouse).
OC  Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX  NCBI_TaxID=10090;
RN  [1]
RP  SEQUENCE FROM N.A.
RA  MEDLINE=92196134; PubMed=1372445;
RX  Kiyokawa H., Busquets X., Powell C.T., Ngo L., Rifkind R.A.,
RA  Marks P.A.;
RT  "Cloning of a D-type cyclin from murine erythroleukemia cells.";
RN  Proc. Natl. Acad. Sci. U.S.A. 89:2444-2447(1992).
RN  [2]
RP  SEQUENCE FROM N.A.
RA  MEDLINE=91235305; PubMed=1827757;
RX  Matsushime H., Roussel M.F., Ashmun R.A., Sherr C.J.;
RT  "Colony-stimulating factor 1 regulates novel cyclins during the G1
RT  phase of the cell cycle.";
RN  Cell 65:701-713(1991).
CC  -1- FUNCTION: ESSENTIAL FOR THE CONTROL OF THE CELL CYCLE AT THE G1/S
CC  (START) TRANSITION.
CC  -1- SUBUNIT: INTERACTS WITH THE CDK4 AND CDK6 PROTEIN KINASES TO FORM
CC  A SERINE/THREONINE KINASE HOLOENZYME COMPLEX. THE CYCLIN SUBUNIT
CC  IMPARTS SUBSTRATE SPECIFICITY TO THE COMPLEX.
CC  -1- SIMILARITY: BELONGS TO THE CYCLIN FAMILY. CYCLIN D SUBFAMILY.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; M83749; AAA37519.1; -
DR  EMBL; M86182; AAA37503.1; -
DR  PIR; B40035; B40035.
DR  PIR; A41984; A41984.
DR  MGD; MGI:88314; Ccnd2.
DR  InterPro; IPR004366; Cyclin_Cterm.
DR  InterPro; IPR004367; Cyclin_C.
DR  Pfam; PF00134; cyclin; 1.
DR  Pfam; PF02984; cyclin_C; 1.
DR  SMART; SM00385; CYCLIN; 1.
DR  PROSITE; PS00292; CYCLINS; 1.
DR  Cyclin; Cell cycle; Cell division; Multigene family.
DR  SEQUENCE 289 AA; 32897 MW; 58F322771DD1DA3D CRC64;

Query Match 61.5%; Score 32; DB 1; Length 289;
Best Local Similarity 60.0%; Pred. No. 29;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10
DB 73 EEVFPPLAMNY 82

RESULT 21
CGD1_BRARE
ID CGD1_BRARE STANDARD; PRT; 291 AA.
AC Q90459;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE  G1/S-specific cyclin D1.
GN  CYCD1.
OS  Brachydanio rerio (zebrafish) (Danio rerio).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC  Cyprinidae; Danio.
OX  NCBI_TaxID=7955;
RN  [1]

```

```

RP SEQUENCE FROM N.A.
RX MEDLINE=96138542; PubMed=8547308;
RA Yarden A., Salomon D., Geiger B.;
RT "Zebrafish cyclin D1 is differentially expressed during early
RL embryogenesis.";
RL Biochim. Biophys. Acta 1264:257-260(1995).
CC -1- FUNCTION: ESSENTIAL FOR THE CONTROL OF THE CELL CYCLE AT THE G1/S
CC (START) TRANSITION.
CC -1- SUBUNIT: INTERACTS WITH THE CDK4 AND CDK6 PROTEIN KINASES TO FORM
CC A SERINE/THREONINE KINASE HOLOENZYME COMPLEX. THE CYCLIN SUBUNIT
CC IMPARTS SUBSTRATE SPECIFICITY TO THE COMPLEX.
CC -1- SIMILARITY: BELONGS TO THE CYCLIN FAMILY. CYCLIN D SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X87581; CAA60885.1;
DR ZFIN; ZDB-GENE-980526-176; cycd1.
DR InterPro; IPR004366; Cyclin.
DR InterPro; IPR004367; Cyclin_Cterm.
DR Pfam; PF00134; cyclin; 1.
DR Pfam; PF02984; cyclin_C; 1.
DR SMART; SM00385; CYCLIN; 1.
DR PROSITE; PS00292; CYCLINS; 1.
DR Cyclin; Cell cycle; Cell division.
KW CYCLIN; Cell cycle; Cell division.
SQ SEQUENCE 291 AA; 33067 MW; FA5274CB1B46D5EF CRC64;

Query Match 61.5%; Score 32; DB 1; Length 291;
Best Local Similarity 60.0%; Pred. No. 29;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10
DB 75 EEVFFPLAMNY 84

RESULT 22
CGDL_XENLA STANDARD; PRT; 291 AA.
AC PS0755;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE G1/S-specific cyclin D1.
GN CCND1.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RA Cockerill M.J., Hunt T.;
RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: ESSENTIAL FOR THE CONTROL OF THE CELL CYCLE AT THE G1/S
CC (START) TRANSITION.
CC -1- SUBUNIT: INTERACTS WITH THE CDK4 AND CDK6 PROTEIN KINASES TO FORM
CC A SERINE/THREONINE KINASE HOLOENZYME COMPLEX. THE CYCLIN SUBUNIT
CC IMPARTS SUBSTRATE SPECIFICITY TO THE COMPLEX.
CC -1- SIMILARITY: BELONGS TO THE CYCLIN FAMILY. CYCLIN D SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----

```

```

CC EMBL; X89475; CAA61664.1;
DR InterPro; IPR004366; Cyclin.
DR InterPro; IPR004367; Cyclin_Cterm.
DR Pfam; PF00134; cyclin; 1.
DR Pfam; PF02984; cyclin_C; 1.
DR SMART; SM00385; CYCLIN; 1.
DR PROSITE; PS00292; CYCLINS; 1.
DR Cyclin; Cell cycle; Cell division; Multigene family.
KW CYCLIN; Cell cycle; Cell division; Multigene family.
SQ SEQUENCE 291 AA; 32953 MW; A4747C5BD1679087 CRC64;

Query Match 61.5%; Score 32; DB 1; Length 291;
Best Local Similarity 60.0%; Pred. No. 29;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10
DB 73 EEVFFPLAMNY 82

RESULT 23
CGD2_CHICK STANDARD; PRT; 291 AA.
AC P49706;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE G1/S-specific cyclin D2.
GN CCND2.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=96144302; PubMed=8566807;
RA Li H., Grenet J., Kidd V.J.;
RT "Structure and gene expression of avian cyclin D2.";
RL Gene 167:341-342(1995).
CC -1- FUNCTION: ESSENTIAL FOR THE CONTROL OF THE CELL CYCLE AT THE G1/S
CC (START) TRANSITION.
CC -1- SUBUNIT: INTERACTS WITH THE CDK4 AND CDK6 PROTEIN KINASES TO FORM
CC A SERINE/THREONINE KINASE HOLOENZYME COMPLEX. THE CYCLIN SUBUNIT
CC IMPARTS SUBSTRATE SPECIFICITY TO THE COMPLEX.
CC -1- SIMILARITY: BELONGS TO THE CYCLIN FAMILY. CYCLIN D SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U28980; AAA96955.1;
DR InterPro; IPR004366; Cyclin.
DR InterPro; IPR004367; Cyclin_Cterm.
DR Pfam; PF00134; cyclin; 1.
DR Pfam; PF02984; cyclin_C; 1.
DR SMART; SM00385; CYCLIN; 1.
DR PROSITE; PS00292; CYCLINS; 1.
DR Cyclin; Cell cycle; Cell division; Multigene family.
KW CYCLIN; Cell cycle; Cell division; Multigene family.
SQ SEQUENCE 291 AA; 33163 MW; 43A7E646AEAF3109 CRC64;

Query Match 61.5%; Score 32; DB 1; Length 291;
Best Local Similarity 60.0%; Pred. No. 29;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10
DB 74 EEVFFPLAMNY 83

```

OC  
OX NCBI\_TaxID=9031;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Li H., Lahti J.M., Kidd V.J.;  
RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: ESSENTIAL FOR THE CONTROL OF THE CELL CYCLE AT THE G1/S  
CC (START) TRANSITION. INTERACTS WITH THE CDK4 AND CDK6 PROTEIN  
CC KINASES (BY SIMILARITY).  
CC -1- SIMILARITY: BELONGS TO THE CYCLIN FAMILY. CYCLIN D SUBFAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: U40844; AAA83271.1; -  
DR InterPro; IPR004366; Cyclin.  
DR InterPro; IPR004367; Cyclin\_Cterm.  
DR Pfam; PF00134; cyclin; 1.  
DR Pfam; PF02984; cyclin\_C; 1.  
DR SMART; SM00385; CYCLIN; 1.  
DR PROSITE; PS00292; CYCLINS; 1.  
KW Cyclin; Cell cycle; Cell division; Multigene family.  
SQ SEQUENCE 292 AA; 33263 MW; 7B543029DB45A67D CRC64;  
  
Query Match 61.5%; Score 32; DB 1; Length 292;  
Best Local Similarity 60.0%; Pred. No. 30;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
  
QY 1 EEVVPXGMSY 10  
Db 75 EEVFPAMNY 84  
  
Search completed: June 10, 2003, 13:40:20  
Job time : 5.5 secs

RESULT 24  
CGD2\_XENLA STANDARD; PRT; 291 AA.  
ID CGD2\_XENLA  
AC P53782;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE G1/S-specific cyclin D2.  
GN CCND2.  
OS Xenopus laevis (African clawed frog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipiloidea; Pipidae;  
OC Xenopodinae; Xenopus.  
OX NCBI\_TaxID=8355;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Cockerill M.J., Hunt T.; to the EMBL/GenBank/DBJ databases.  
RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97380591; PubMed=9237366;  
RA Taieb F., Jessus C.;  
RT "Xenopus cyclin D2: cloning and expression in oocytes and during  
RT early development.";  
RL Biol. Cell 88:99-111(1996).  
CC -1- FUNCTION: ESSENTIAL FOR THE CONTROL OF THE CELL CYCLE AT THE G1/S  
CC (START) TRANSITION.  
CC -1- SUBUNIT: INTERACTS WITH THE CDK4 AND CDK6 PROTEIN KINASES TO FORM  
CC A SERINE/THREONINE KINASE HOLOENZYME COMPLEX. THE CYCLIN SUBUNIT  
CC IMPARTS SUBSTRATE SPECIFICITY TO THE COMPLEX.  
CC -1- SIMILARITY: BELONGS TO THE CYCLIN FAMILY. CYCLIN D SUBFAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: X89476; CAA61665.1; -  
DR EMBL: X83503; CAA58493.1; -  
DR InterPro; IPR004366; Cyclin.  
DR InterPro; IPR004367; Cyclin\_Cterm.  
DR Pfam; PF00134; cyclin; 1.  
DR Pfam; PF02984; cyclin\_C; 1.  
DR SMART; SM00385; CYCLIN; 1.  
DR PROSITE; PS00292; CYCLINS; 1.  
KW Cyclin; Cell cycle; Cell division; Multigene family.  
SQ SEQUENCE 291 AA; 32959 MW; 9A290F04F1531E89 CRC64;  
  
Query Match 61.5%; Score 32; DB 1; Length 291;  
Best Local Similarity 60.0%; Pred. No. 29;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
  
QY 1 EEVVPXGMSY 10  
Db 74 EEVFPAMNY 83

RESULT 25  
CGD1\_CHICK STANDARD; PRT; 292 AA.  
ID CGD1\_CHICK  
AC P55169;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE G1/S-specific cyclin D1.  
GN CCND1.  
OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2003, 13:30:25 ; Search time 25.7857 Seconds  
(without alignments)  
87.898 Million cell updates/sec

Title: US-09-909-164-9  
Perfect score: 52  
Sequence: 1 EEWVPGMSYS 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues 671580  
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 75 summaries

Database : SPTREMBL\_21:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phase:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_rvirus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	38	73.1	387	16 Q98FX1	Q98fx1 rhizobium 1
2	38	73.1	1063	16 Q9RG86	Q9rg86 fusobacteri
3	38	73.1	3472	1 Q74056	Q74056 cenarchaeum
4	37	71.2	840	3 Q9URY8	Q9ury8 schizosacch
5	36	69.2	471	11 Q8RL26	Q8rl26 mus musculu
6	36	69.2	484	11 Q8VD18	Q8vd18 mus musculu
7	35	67.3	225	10 Q40129	Q40129 lycopersico
8	35	67.3	425	5 Q9XVK4	Q9xvk4 caenorhabdi
9	35	67.3	556	4 Q43733	Q43733 homo sapien
10	35	67.3	583	5 Q9BHA5	Q9bha5 plasmodium
11	35	67.3	583	5 Q9BHA3	Q9bha3 plasmodium
12	35	67.3	670	11 Q01487	Q01487 rattus norv
13	35	67.3	749	16 Q9PDM6	Q9pdm6 xylella fas
14	35	67.3	1902	4 Q14122	Q14122 homo sapien
15	34	65.4	156	3 Q12479	Q12479 saccharomyc
16	34	65.4	219	17 Q971S2	Q971s2 sulfolobus

17	34	65.4	252	17 Q28342	Q28342 archaeoglob
18	34	65.4	250	4 Q96MU1	Q96mu1 homo sapien
19	34	65.4	387	16 Q92MD6	Q92md6 rhizobium m
20	34	65.4	541	16 Q98BP5	Q98bp5 rhizobium l
21	34	65.4	544	16 Q9PQD2	Q9pqd2 ureaplasma
22	34	65.4	842	3 Q9URR4	Q9urr4 penicillium
23	34	65.4	1049	16 Q8XT05	Q8xt05 ralstonia s
24	34	65.4	1499	4 Q96914	Q96914 homo sapien
25	33	63.5	143	17 Q8TX62	Q8tx62 methanopyru
26	33	63.5	162	11 Q9CXQ4	Q9cxq4 mus musculu
27	33	63.5	165	17 Q28330	Q28330 archaeoglob
28	33	63.5	193	2 Q8VUA8	Q8vua8 lactococcus
29	33	63.5	209	16 Q8RE56	Q8re56 fusobacteri
30	33	63.5	284	16 P74187	P74187 synecocyst
31	33	63.5	298	10 Q9M3C0	Q9m3c0 arabidopsis
32	33	63.5	326	12 Q9Q9Q9	Q9q9q9 soil-borne
33	33	63.5	326	12 Q9Q9Q5	Q9q9q5 soil-borne
34	33	63.5	326	12 Q9QCE7	Q9qce7 soil-borne
35	33	63.5	326	12 Q9DJG4	Q9djg4 soil-borne
36	33	63.5	326	12 Q91DN1	Q91dn1 soil-borne
37	33	63.5	326	12 Q9Q9Q7	Q9q9q7 soil-borne
38	33	63.5	327	12 Q06360	Q06360 soil-borne
39	33	63.5	332	10 Q9FNL4	Q9fnl4 arabidopsis
40	33	63.5	368	16 Q9X0U3	Q9x0u3 thermotoga
41	33	63.5	393	5 Q9V914	Q9v914 drosophila
42	33	63.5	479	4 Q96CS0	Q96cs0 homo sapien
43	33	63.5	548	11 Q9D2X9	Q9d2x9 mus musculu
44	33	63.5	584	16 Q8R8K6	Q8r8k6 thermoanaer
45	33	63.5	648	4 Q96MB2	Q96mb2 homo sapien
46	33	63.5	653	16 Q9KVE3	Q9kve3 vibrio chol
47	33	63.5	676	5 Q9VA55	Q9va55 drosophila
48	33	63.5	676	5 Q8T8Z7	Q8t8z7 drosophila
49	33	63.5	678	12 Q9E1X6	Q9e1x6 cercopithe
50	33	63.5	746	3 Q9URR3	Q9urr3 penicillium
51	33	63.5	791	4 Q9H2K5	Q9h2k5 homo sapien
52	33	63.5	793	4 Q9H2K6	Q9h2k6 homo sapien
53	33	63.5	844	11 Q922D4	Q922d4 mus musculu
54	33	63.5	1028	16 Q8YJ11	Q8yj11 brucella me
55	33	63.5	1152	16 Q9CC95	Q9cc95 mycobacteri
56	33	63.5	1305	5 Q9V7C7	Q9v7c7 drosophila
57	33	63.5	1394	4 Q8TD95	Q8td95 homo sapien
58	33	63.5	1442	17 Q96YH5	Q96yh5 sulfolobus
59	33	63.5	1548	10 Q85531	Q85531 arabidopsis
60	33	63.5	1713	11 Q88349	Q88349 mus musculu
61	32	61.5	84	16 Q97DE7	Q97de7 clostridium
62	32	61.5	103	11 Q9D0H9	Q9d0h9 mus musculu
63	32	61.5	108	1 Q9UX33	Q9ux33 sulfolobus
64	32	61.5	153	13 Q79919	P79919 xenopus lae
65	32	61.5	156	11 Q9D8L9	Q9d8l9 mus musculu
66	32	61.5	174	10 Q9M3T4	Q9m3t4 betula verr
67	32	61.5	190	13 Q57481	Q57481 stizostedio
68	32	61.5	191	11 Q99NB4	Q99nb4 rattus norv
69	32	61.5	200	17 Q97CD0	Q97cd0 thermoplasm
70	32	61.5	207	2 Q47284	Q47284 escherichia
71	32	61.5	234	2 Q32330	Q32330 clostridium
72	32	61.5	236	10 Q9SXF1	Q9sxf1 arabidopsis
73	32	61.5	240	11 Q9DB09	Q9db09 mus musculu
74	32	61.5	243	12 Q91EW1	Q91ew1 cydia pomon
75	32	61.5	279	17 Q9Y8Z4	Q9y8z4 aeropyrum p

# ALIGNMENTS

RESULT 1  
Q98FX1 PRELIMINARY; PRT; 387 AA.  
ID Q98FX1  
AC Q98FX1;  
DT 01-OCT-2001 (TREMBLrel. 18, Created)  
DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)  
DE 01-OCT-2001 (TREMBLrel. 18, Last annotation update)  
DE Hippurate hydrolase.  
GN MLR3583.

```

OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Idegawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki M., Shimpō S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti."
RL DNA Res. 7:331-338(2000).
DR EMBL: AP003002; BAB50445.1; -
DR InterPro: IPR002933; Peptidase_M20.
DR Pfam: PF01546; Peptidase_M20; 1.
DR Hydrolase; Complete proteome.
SQ SEQUENCE 387 AA; 41180 MW; 131BFF8E64306829 CRC64;

Query Match 73.1%; Score 38; DB 16; Length 387;
Best Local Similarity 60.0%; Pred. No. 9.8;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10
Db 367 DEAIPIHGMSY 376

RESULT 2
Q8RG86 PRELIMINARY; PRT; 1063 AA.
ID Q8RG86
AC Q8RG86;
DT 01-JUN-2002 (Tremblrel. 21, Created)
DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Carbamoyl-phosphate synthase large chain (EC 6.3.5.5).
GN FN0422.
OS Fusobacterium nucleatum (subsp. nucleatum).
OC Bacteria; Fusobacteriia; Fusobacterium.
OX NCBI_TaxID=76856;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 25586;
RX MEDLINE=21886394; PubMed=11889109;
RA Kapatal V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A.,
RA Bhattacharyya A., Bartman A., Gardner W., Grechkin G., Zhu L.,
RA Vasileva O., Chu L., Kogan Y., Chaga O., Goltzman E., Bernal A.,
RA Larsen N., D'Souza M., Walunas T., Pusck G., Haselkorn R.,
RA Fonstein M., Kyrpides N., Overbeek R.;
RT "Genome sequence and analysis of the oral bacterium Fusobacterium
RT nucleatum strain ATCC 25586."
RL J. Bacteriol. 184:2005-2018(2002).
DR EMBL: AE010554; AAL94625.1; -
DR Ligase; Complete proteome.
SQ SEQUENCE 1063 AA; 118008 MW; 39700E10B7CCE411 CRC64;

Query Match 73.1%; Score 38; DB 16; Length 1063;
Best Local Similarity 60.0%; Pred. No. 29;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 EEVVPXGMSY 11
Db 195 EIVPGLNYS 204

RESULT 3
O74056 PRELIMINARY; PRT; 3472 AA.
ID O74056
AC O74056;
DT 01-NOV-1998 (Tremblrel. 08, Created)

```

```

DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Hypothetical 367.1 kDa protein.
OS Cenarchaeum symbiosum.
OC Archaea; Crenarchaeota; Cenarchaeum.
OX NCBI_TaxID=46770;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=B.
RX MEDLINE=98422450; PubMed=9748430;
RA Schleper C., DeLong E.F., Preston C.M., Feldman R.A., Wu K.Y.,
RA Swanson R.V.;
RT "Genomic analysis reveals chromosomal variation in natural populations
RT of the uncultured psychrophilic archaeon Cenarchaeum symbiosum."
RL J. Bacteriol. 180:5003-5009(1998).
CC 1- SIMILARITY: CONTAINS 5 WD REPEATS (TRP-ASP DOMAINS).
DR EMBL: AF083072; AAC62699.1; -
DR InterPro: IPR00515; BPD.transp.
DR Pfam: PF00400; WD40; 4.
DR SMART; SM00320; WD40; 2.
DR PROSITE; PS00402; BPD_TRANS_INN_MEMBR; UNKNOWN_1.
KW Hypothetical protein; Repeat; WD repeat.
SQ SEQUENCE 3472 AA; 367058 MW; 37F80707030F9355 CRC64;

Query Match 73.1%; Score 38; DB 1; Length 3472;
Best Local Similarity 54.5%; Pred. No. 1e+02;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 11
Db 2294 EDVIPRGISFS 2304

RESULT 4
Q9URY8 PRELIMINARY; PRT; 840 AA.
ID Q9URY8
AC Q9URY8;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Probable sulfate permease.
GN SPAC869.05C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972H-;
RA Hunt C., Aves S., McDougall R.C., Rajandream M.A., Barrell B.G.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL132779; CAB60015.1; -
DR InterPro: IPR002645; STAS.
DR InterPro: IPR001902; Sulfate_transp.
DR Pfam: PF01740; STAS; 1.
DR Pfam: PF00916; Sulfate_transp; 1.
DR TIGRFS: TIGR00815; sulp; 1.
SQ SEQUENCE 840 AA; 93517 MW; ED4833E162B69077 CRC64;

Query Match 71.2%; Score 37; DB 3; Length 840;
Best Local Similarity 77.8%; Pred. No. 37;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMSY 11
Db 135 VVPQGMYSY 143

RESULT 5
Q8R126 PRELIMINARY; PRT; 471 AA.
ID Q8R126

```



Q8R126;  
 01-JUN-2002 (TrEMBLrel. 21, Created)  
 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 Hypothetical 54.5 kDa protein (Fragment).  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=LIVER;  
 RA Strausberg R.;  
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
 KW Hypothetical protein.  
 FT NON\_TER 1  
 SQ SEQUENCE 471 AA; 54506 MW; E0DA685C374A9760 CRC64;  
 Query Match 69.2%; Score 36; DB 11; Length 471;  
 Best Local Similarity 60.0%; Pred. No. 32;  
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
 Qy 2 EVVPXGMSYS 11  
 Db 226 EVIPAGASYN 235  
 RESULT 6  
 Q8VD18  
 ID Q8VD18 PRELIMINARY; PRT; 484 AA.  
 AC Q8VD18;  
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE Similar to glioma tumor suppressor candidate region gene 2.  
 GN AWS36441.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=SALIVARY GLAND;  
 RA Strausberg R.;  
 RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC017637; AAH17637.1;  
 DR MGD; MGI:2138595; AW536441.  
 SQ SEQUENCE 484 AA; 55835 MW; BBB45F3B4BE02A36 CRC64;  
 Query Match 69.2%; Score 36; DB 11; Length 484;  
 Best Local Similarity 60.0%; Pred. No. 33;  
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
 Qy 2 EVVPXGMSYS 11  
 Db 239 EVIPAGASYN 248  
 RESULT 7  
 Q40129  
 ID Q40129 PRELIMINARY; PRT; 225 AA.  
 AC Q40129;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 DE Hypothetical 25.2 kDa protein precursor.  
 OS Lycopersicon esculentum (Tomato).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.  
 OX NCBI\_TaxID=4081;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RC STRAIN=VF36; TISSUE=PISTIL;  
 RX MEDLINE=95375233; Pubmed=7647301;  
 RA Milligan S.B.; Gasser C.S.;  
 RT "Nature and regulation of pistil-expressed genes in tomato.";  
 RL Plant Mol. Biol. 28:691-711(1995).  
 DR EMBL; U20592; AAA80497.1;  
 DR InterPro; IPR002160; Kunitz\_legume.  
 DR Pfam; PF00197; Kunitz\_legume; 1.  
 DR ProDom; PD000891; Kunitz\_legume; 1.  
 DR SMART; SM00452; STI; 1.  
 DR PROSITE; PS00283; SOYBEAN\_KUNITZ; UNKNOWN\_1.  
 KW Hypothetical protein; Signal.  
 FT SIGNAL 1 20 POTENTIAL.  
 CHAIN 21 225 UNKNOWN.  
 SQ SEQUENCE 225 AA; 25188 MW; 1074C261D20CFDAD CRC64;  
 Query Match 67.3%; Score 35; DB 10; Length 225;  
 Best Local Similarity 54.5%; Pred. No. 24;  
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
 Qy 1 EVVPXGMSYS 11  
 Db 32 DEVVPNGKTYA 42  
 RESULT 8  
 Q9XVK4  
 ID Q9XVK4 PRELIMINARY; PRT; 425 AA.  
 AC Q9XVK4;  
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 DE R10D12.10 protein.  
 GN R10D12.10.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Percy C.M.;  
 RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99069613; Pubmed=9851916;  
 RA none;  
 RT "Genome sequence of the nematode C. elegans: A platform for investigating biology."  
 RL Science 282:2012-2018(1998).  
 DR EMBL; Z81109; CAB03241.1;  
 DR InterPro; IPR000719; Euk\_pkinase.  
 DR Pfam; PF00069; pkinase; 1.  
 DR ProDom; PD000001; Euk\_pkinase; 1.  
 DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
 KW ATP-binding; transferase.  
 SQ SEQUENCE 425 AA; 49410 MW; 5D96E29B08C8E9D6 CRC64;  
 Query Match 67.3%; Score 35; DB 5; Length 425;  
 Best Local Similarity 50.0%; Pred. No. 48;  
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
 Qy 1 EVVPXGMSYS 10  
 Db 335 EQIVPGGLQY 344  
 RESULT 9  
 O43733  
 ID O43733 PRELIMINARY; PRT; 556 AA.  
 AC O43733;  
 DT 01-JUN-1998 (TrEMBLrel. 06, Created)  
 DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)

DT 01-JUN-1998 (TREMBlrel. 06, Last annotation update)  
 DE DNA binding protein (Fragment).  
 GN DJ451B15.2.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Tubby B.;  
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; Z98050; CAB10847.1; -  
 FT NON\_TER 1  
 SQ SEQUENCE 556 AA; 59059 MW; ECB00E4033FB2528 CRC64;

Query Match 67.3%; Score 35; DB 4; Length 556;  
 Best Local Similarity 66.7%; Pred. No. 64;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 QY 3 VVPXGMSYS 11  
 Db 244 VVPAGLTYS 252  
 ||| |::|||

RESULT 10  
 Q9BH85 PRELIMINARY; PRT; 583 AA.  
 ID Q9BH85  
 AC Q9BH85  
 DT 01-JUN-2001 (TREMBlrel. 17, Created)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE Choline transporter.  
 GN SCT1.  
 OS Plasmodium falciparum.  
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
 OX NCBI\_TaxID=5833;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Ben Mamoun C., Gluzman I.Y., Goldberg D.E.;  
 RT "Plasmodium falciparum choline transporter (pfSCT1) gene."  
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY007372; AAK14816.1; -  
 DR EMBL; AY007375; AAG17947.1; -  
 DR InterPro; IPR002123; Acyltransferase.  
 DR Pfam; PF01553; Acyltransferase; 1.  
 SQ SEQUENCE 583 AA; 66917 MW; 2B2BFAE3E395E049 CRC64;

Query Match

Best Local Similarity 67.3%; Score 35; DB 5; Length 583;  
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
 QY 3 VVPXGMSYS 11  
 Db 227 IIPVGLSYS 235  
 ||| |::|||

RESULT 11  
 Q9BH83 PRELIMINARY; PRT; 583 AA.  
 ID Q9BH83  
 AC Q9BH83  
 DT 01-JUN-2001 (TREMBlrel. 17, Created)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE Choline transporter.  
 GN SCT1.  
 OS Plasmodium falciparum.  
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
 OX NCBI\_TaxID=5833;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Ben Mamoun C., Gluzman I.Y., Goldberg D.E.;  
 RT "Plasmodium falciparum choline transporter (pfSCT1) gene."  
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY007372; AAK14816.1; -  
 DR EMBL; AY007375; AAG17947.1; -  
 DR InterPro; IPR002123; Acyltransferase.  
 DR Pfam; PF01553; Acyltransferase; 1.  
 SQ SEQUENCE 583 AA; 66917 MW; 2B2BFAE3E395E049 CRC64;

Query Match

Best Local Similarity 67.3%; Score 35; DB 5; Length 583;  
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
 QY 3 VVPXGMSYS 11  
 Db 227 IIPVGLSYS 235  
 ||| |::|||

RESULT 11  
 Q9BH83 PRELIMINARY; PRT; 583 AA.  
 ID Q9BH83  
 AC Q9BH83  
 DT 01-JUN-2001 (TREMBlrel. 17, Created)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE Choline transporter.  
 GN SCT1.  
 OS Plasmodium falciparum.  
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
 OX NCBI\_TaxID=5833;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Ben Mamoun C., Gluzman I.Y., Goldberg D.E.;  
 RT "Plasmodium falciparum choline transporter (pfSCT1) gene."  
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.

Query Match

Best Local Similarity 67.3%; Score 35; DB 5; Length 583;  
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
 QY 3 VVPXGMSYS 11  
 Db 227 IIPVGLSYS 235  
 ||| |::|||

RESULT 11  
 Q9BH83 PRELIMINARY; PRT; 583 AA.  
 ID Q9BH83  
 AC Q9BH83  
 DT 01-JUN-2001 (TREMBlrel. 17, Created)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE Choline transporter.  
 GN SCT1.  
 OS Plasmodium falciparum.  
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
 OX NCBI\_TaxID=5833;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Ben Mamoun C., Gluzman I.Y., Goldberg D.E.;  
 RT "Plasmodium falciparum choline transporter (pfSCT1) gene."  
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.

Query Match

Best Local Similarity 66.7%; Pred. No. 78;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 QY 3 VVPXGMSYS 11  
 Db 376 VVPAGLTYS 384  
 ||| |::|||

RESULT 13

DR EMBL; AY007374; AAK14818.1; -  
 DR EMBL; AY007373; AAK14817.1; -  
 DR InterPro; IPR002123; Acyltransferase.  
 DR Pfam; PF01553; Acyltransferase; 1.  
 SQ SEQUENCE 583 AA; 66918 MW; 4DF83D7530527474 CRC64;

Query Match 67.3%; Score 35; DB 5; Length 583;  
 Best Local Similarity 55.6%; Pred. No. 67;  
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
 QY 3 VVPXGMSYS 11  
 Db 227 IIPVGLSYS 235  
 ||| |::|||

RESULT 12  
 Q01487 PRELIMINARY; PRT; 670 AA.  
 ID Q01487  
 AC Q01487  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
 DE DNA-binding protein AT-BP2 (ALPHA1-antitrypsin promoter binding protein 2) (Fragment).  
 DE protein 2) (Fragment).  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-SPRAGUE-DAWLEY; TISSUE=THYROID;  
 RX MEDLINE-91187610; PubMed-1901405;  
 RA Mitchelmore C., Traboni C., Cortese R.;  
 RT "Isolation of two cDNAs encoding zinc finger proteins which bind to the alpha 1-antitrypsin promoter and to the major histocompatibility complex class I enhancer."  
 RT Nucleic Acids Res. 19:141-147(1991).  
 RL CC  
 CC -!- FUNCTION: BINDS TO THE ALPHA1-ANTITRYPSIN PROMOTER, TO THE KAPPA IMMUNOGLOBULIN GENE ENHANCER, AND TO THE MAJOR HISTOCOMPATIBILITY COMPLEX CLASS I ENHANCER; IT PLAYS A ROLE AS TRANSCRIPTIONAL REGULATOR. CONCERNING THE ALPHA1-ANTITRYPSIN EXPRESSION IT MAY ACT AS A NEGATIVE REGULATOR BY INTERFERING AT THE ALPHA1-ANTITRYPSIN PROMOTOR WITH ANOTHER NUCLEAR PROTEIN, CALLED LF-B.  
 CC -!- SUBCELLULAR LOCATION: NUCLEAR.  
 CC -!- MISCELLANEOUS: ACIDIC AMINO ACID REGION MAY BE INVOLVED IN TRANSCRIPTIONAL ACTIVATION.  
 CC -!- SIMILARITY: HIGHLY SIMILAR TO AT-BP1 AND TO THE PRDII-BF1 FACTOR.  
 DR EMBL; X54250; CAA38151.1; -  
 DR HSP; P15822; IBBO.  
 DR InterPro; IPR000822; Znf\_C2H2.  
 DR Pfam; PF00006; zf-C2H2; 2.  
 DR SMART; SM00355; Znf\_C2H2; 2.  
 DR PROSITE; PS00028; ZINC\_FINGER\_C2H2\_1; 2.  
 DR PROSITE; PS00157; ZINC\_FINGER\_C2H2\_2; 2.  
 KW Zinc-finger; Nuclear protein; DNA-binding; Transcription regulation;  
 KW Metal-binding; Multigene family.  
 FT NON\_TER 1  
 FT DOMAIN 54 104 ZINC-FINGERS.  
 FT DOMAIN 140 160 ACIDIC.  
 FT ZN\_FING 54 74 C(2)H(2) CLASS.  
 FT ZN\_FING 82 104 C(2)H(2) CLASS.  
 SQ SEQUENCE 670 AA; 71233 MW; CDD2324152590C17 CRC64;

Query Match 67.3%; Score 35; DB 11; Length 670;  
 Best Local Similarity 66.7%; Pred. No. 78;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 QY 3 VVPXGMSYS 11  
 Db 376 VVPAGLTYS 384  
 ||| |::|||

RESULT 13

Q9PDM6 Q9PDM6 PRELIMINARY; PRT; 749 AA.  
 ID Q9PDM6  
 AC Q9PDM6  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE Topoisomerase IV subunit.  
 GN XF1353.  
 OS Xylella fastidiosa.  
 OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;  
 OC Xylella.  
 OC NCBI\_TaxID=2371;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=9A5C;  
 RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,  
 RA Alvarenga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,  
 RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,  
 RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carer H.,  
 RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,  
 RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,  
 RA Facinani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,  
 RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furian L.R.,  
 RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,  
 RA Ho P.L., Hohelsel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,  
 RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite D.C.C.,  
 RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,  
 RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,  
 RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,  
 RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,  
 RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,  
 RA Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,  
 RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,  
 RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,  
 RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,  
 RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,  
 RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,  
 RA da Silva J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,  
 RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsubako M.H.,  
 RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,  
 RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;  
 RT "The genome sequence of the plant pathogen Xylella fastidiosa";  
 RL Nature 406:151-159(2000).  
 DR EMBL; AF003967; XAF84162.1;  
 DR HSSP; P09097; IAB4.  
 DR InterPro; IPR002205; DNA\_topoisomIV.  
 DR Pfam; PF00521; DNA\_topoisomIV; 1.  
 DR ProDom; PD000742; DNA\_topoisomIV; 1.  
 DR SMART; SM00434; TOP4c; 1.  
 DR TIGRFAMs; TIGR01062; parC\_Gneg; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 749 AA; 83344 MW; 4DCD10F480EE0257 CRC64;  
 Query Match 67.3%; Score 35; DB 16; Length 749;  
 Best Local Similarity 77.8%; Pred. No. 88;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 2 EVVPGMSY 10  
 Db 526. EVDPSGMSY 534  
 RESULT 14  
 Q14122 Q14122 PRELIMINARY; PRT; 1902 AA.  
 ID Q14122  
 AC Q14122  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 DE DNA-binding protein (Mbp-1) (Fragment).  
 OS Homo sapiens (human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Euthera; Primates; Catarrhini; Hominoidea; Homo.  
 OC NCBI\_TaxID=9606;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=90205817; PubMed=2108316;  
 RA Baldwin A.S., LeClair K.P., Singh H., Sharp P.A.;  
 RT "A large protein containing zinc finger domains binds to related  
 RT sequence elements in the enhancers of the class I major  
 RT histocompatibility complex and kappa immunoglobulin genes.";  
 RL Mol. Cell. Biol. 10:1406-1414(1990).  
 DR EMBL; M32019; AAA17534.1;  
 DR HSSP; P15822; 1BBO.  
 DR InterPro; IPR000822; Znf\_C2H2.  
 DR Pfam; PF00096; zf-C2H2; 3.  
 DR SMART; SM00355; Znf\_C2H2; 2.  
 DR PROSITE; PS00028; ZINC\_FINGER\_C2H2\_1; 2.  
 DR PROSITE; PS00157; ZINC\_FINGER\_C2H2\_2; 2.  
 KW DNA-binding; Metal-binding; Zinc-finger.  
 FT NON\_TER 1  
 SQ SEQUENCE 1902 AA; 207457 MW; 35BBF0D961E5ED20 CRC64;  
 Query Match 67.3%; Score 35; DB 4; Length 1902;  
 Best Local Similarity 66.7%; Pred. No. 2.4e+02;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 QY 3 VVPXGMSYS 11  
 Db 1590 VVPAGLTYS 1598  
 RESULT 15  
 Q12479 Q12479 PRELIMINARY; PRT; 156 AA.  
 ID Q12479  
 AC Q12479  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE ORF YOR013W.  
 GN YOR013W.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
 OC NCBI\_TaxID=4932;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RC De haan M., Grivell L.A., Maarse A.C.;  
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
 [2]  
 RP SEQUENCE FROM N.A.  
 RA MIPS;  
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
 [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=FY1679;  
 RA De haan M., Maarse A.C., Grivell L.A.;  
 RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.  
 [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=FY1679;  
 RA Dumont M.E., Schlichter J.B., Cardillo T.S., Hayes M.K., Bethlenny G.,  
 RA Sherman F.;  
 RT "CYC2 encodes a factor involved in mitochondrial import of yeast  
 RT cytochrome c.";  
 RL Mol. Cell. Biol. 13:6442-6451(1993).  
 [5]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=FY1679;  
 RA MEDLINE=94169519; PubMed=7764548;  
 RA Lee Y.S., Shimizu J., Yoda K., Yamasaki M.;  
 RT "Molecular cloning of a gene, DHS1, which complements a drug-  
 RT hypersensitive mutation of the yeast Saccharomyces cerevisiae";  
 RL Biosci. Biotechnol. Biochem. 58:391-395(1994).

```

DR EMBL; Z74920; CAA99201.1; -
DR EMBL; X87331; CAA60762.1; -
DR SGD; S0005539; YOR013W;
SQ SEQUENCE 156 AA; 17881 MW; 380442B74C272B41 CRC64;

Query Match
Best Local Similarity 65.4%; Score 34; DB 3; Length 156;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 EVVPXGMSY 10
   ||:| | | |
Db 50 EVNPLGMDY 58

RESULT 16
Q971S2 PRELIMINARY; PRT; 219 AA.
AC Q971S2;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DE Putative ribose 5-phosphate isomerase.
GN ST1302.
OS Sulfolobus tokodaii.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.
OX NCBI_TaxID=111955;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-JCM 10545 / 7;
RX PubMed=11572479;
RA Kawarabayashi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,
RA Sekine M., Baba S.-I., Anka A., Kosugi H., Hosoyama A., Fukui S.,
RA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,
RA Yoshizawa T., Tanaka T., Kudo H., Yamazaki J., Kishida N., Oguchi A.,
RA Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A.,
RA Oshima T., Kikuchi H.;
RT "Complete genome sequence of an aerobic thermoacidophilic
RT Crenarchaeon, Sulfolobus tokodaii strain7.";
RL DNA Res. 8:123-140(2001).
DR EMBL; AP000985; BAB66348.1; -
DR InterPro; IPR004788; RpiA.
DR ProDom; PD005813; RpiA; 1.
KW Isomerase; Hypothetical protein; Complete proteome.
SQ SEQUENCE 219 AA; 24541 MW; A4E9A3F2C4006D90 CRC64;

Query Match
Best Local Similarity 65.4%; Score 34; DB 17; Length 219;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 EVVPXGMSY 10
   ||| | | |
Db 131 EVVPVGAVY 139

RESULT 17
O28342 PRELIMINARY; PRT; 252 AA.
AC O28342;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DE Cell division inhibitor (MIND-2).
GN AF1937.
OS Archaeoglobus fulgidus.
OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
OC Archaeoglobaceae; Archaeoglobus.
OX NCBI_TaxID=2234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=98049343; PubMed=9389475;
RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,

```

DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)  
 DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)  
 DE PUTATIVE HIPURATE HYDROLASE PROTEIN (EC 3.5.1.32).  
 GN HIPOL OR R02690 OR SMC00682.  
 OS Rhizobium meliloti (Sinorhizobium meliloti).  
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
 OC Rhizobiaceae; Sinorhizobium.  
 OX NCBI\_TaxID=382;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=1021;  
 RX MEDLINE=21396507; PubMed=11481430;  
 RA Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,  
 RA Bolstad P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,  
 RA Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masuy D.,  
 RA Pohl T., Portetelle D., Puehler A., Purnelle B., Ramsperger U.,  
 RA Renard C., Thebault P., Vandenbol M., Weidner S., Galibert F.,  
 RT "Analysis of the chromosome sequence of the legume symbiont  
 RT Sinorhizobium meliloti strain 1021.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).  
 DR EMBL; AL591791; CAC47269.1;  
 DR InterPro; IPR002933; Peptidase\_M20.  
 DR Pfam; PF01546; Peptidase\_M20; 1.  
 KW Hydrolase; Complete proteome.  
 SQ SEQUENCE 387 AA; 41074 MW; 342763088907A6E3 CRC64;

Query Match 65.4%; Score 34; DB 16; Length 387;  
 Best Local Similarity 50.0%; Pred. No. 70;  
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10  
 DB 367 DEAIHPHGYSY 376

RESULT 20  
 Q98BP5 PRELIMINARY; PRT; 541 AA.  
 AC Q98BP5;  
 DT 01-OCT-2001 (TRENBLrel. 18, Created)  
 DT 01-OCT-2001 (TRENBLrel. 18, Last sequence update)  
 DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)  
 DE Probable DNA ligase.  
 GN MLL3481.  
 OS Rhizobium loti (Mesorhizobium loti).  
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
 OC Phyllobacteriaceae; Mesorhizobium.  
 OX NCBI\_TaxID=381;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MAFF303099;  
 RX MEDLINE=21082930; PubMed=11214968;  
 RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,  
 RA Watanabe A., Iidesawa K., Ishikawa A., Kawashima K., Kimura T.,  
 RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,  
 RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,  
 RA Takeuchi C., Yamada M., Tabata S.,  
 RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium  
 RT Mesorhizobium loti.";  
 RL DNA Res. 7:331-338(2000).  
 DR EMBL; AP003006; BAB51927.1;  
 DR InterPro; IPR000977; DNA\_ligase.  
 DR Pfam; PF01068; DNA\_ligase; 1.  
 DR PROSITE; PS00697; DNA\_LIGASE\_A1; UNKNOWN\_1.  
 DR PROSITE; PS0160; DNA\_LIGASE\_A3; 1.  
 KW Ligase; Complete proteome.  
 SQ SEQUENCE 541 AA; 60645 MW; 2E7EF705453F2BF8 CRC64;

Query Match 65.4%; Score 34; DB 16; Length 541;  
 Best Local Similarity 60.0%; Pred. No. 1e+02;  
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10

Db 445 EELVPVKAY 454

RESULT 21  
 Q9PQD2 PRELIMINARY; PRT; 544 AA.  
 AC Q9PQD2;  
 DT 01-OCT-2000 (TRENBLrel. 15, Created)  
 DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)  
 DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)  
 DE Putative ABC substrate-binding protein-iron.  
 GN ABCSBP-5 OR UU359.  
 OS Ureaplasma parvum (Ureaplasma urealyticum biotype 1).  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;  
 OC Mycoplasmataceae; Ureaplasma.  
 OX NCBI\_TaxID=134821;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=SEROVAR 3;  
 RX MEDLINE=20500219; PubMed=11048724;  
 RA Glass J.I., Lefkowitz E.J., Glass J.S., Heiner C.R., Chen E.Y.,  
 RA Cassell G.H.;  
 RT "The complete sequence of the mucosal pathogen Ureaplasma  
 RT urealyticum.";  
 RL Nature 407:757-762(2000).  
 DR EMBL; AE002133; AAF30768.1;  
 KW Complete proteome.  
 SQ SEQUENCE 544 AA; 61291 MW; CF8756202A389C00 CRC64;

Query Match 65.4%; Score 34; DB 16; Length 544;  
 Best Local Similarity 70.0%; Pred. No. 1e+02;  
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10  
 DB 135 EEVVPYHLSY 144

RESULT 22  
 Q9URR4 PRELIMINARY; PRT; 842 AA.  
 AC Q9URR4;  
 DT 01-MAY-2000 (TRENBLrel. 13, Created)  
 DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)  
 DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)  
 DE Sulfate permease SufB.  
 GN SUTB.  
 OS Penicillium chrysogenum.  
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;  
 OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Penicillium.  
 OX NCBI\_TaxID=5076;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Q176;  
 RX MEDLINE=20042342; PubMed=10572125;  
 RA Van de Kamp M., Pizzini E., Vos A., Van der Lende T.R.,  
 RA Schuur T.A., Newbert R.W., Turner G., Konings W.N., Driessen A.J.M.;  
 RT "Sulfate transport in Penicillium chrysogenum; Cloning and  
 RT Characterization of the sutA and sutB Genes.";  
 RL J. Bacteriol. 181:7228-7234(1999).  
 DR EMBL; AF163974; AAF14539.1;  
 DR InterPro; IPR002645; STAS.  
 DR InterPro; IPR001902; Sulfate\_transp.  
 DR Pfam; PF01740; STAS; 1.  
 DR Pfam; PF00916; Sulfate\_transp; 1.  
 DR TIGRFAMs; TIGR00815; sulp; 1.  
 DR PROSITE; PS01130; SULFATE\_TRANSP; UNKNOWN\_1.  
 SQ SEQUENCE 842 AA; 91865 MW; 839A55486E733D15 CRC64;

Query Match 65.4%; Score 34; DB 3; Length 842;  
 Best Local Similarity 66.7%; Pred. No. 1.6e+02;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10

```

QY      3 VVPXGMSYS 11
Db      111 VVPOGMAYA 119

RESULT 23
Q8XT05 PRELIMINARY; PRT; 1049 AA.
ID Q8XT05
AC Q8XT05;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DE Probable multidrug efflux system transmembrane protein.
GN MEXD OR RSP0312 OR RS05457.
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OC Plasmid megaplasmid.
OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
OC Ralstonia.
ON NCBI_TaxID=305;
RX STRAIN-GM11000;
MEDLINE=21681879; PubMed=11823852;
RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
RA Arlat M., Billault A., Brottier P., Camus J.C., Cattolico L.,
RA Chandler M., Choise N., Claudel-Renard C., Cunha S., Demange N.,
RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schiex T.,
RA Siquier P., Thebaud P., Whalen M., Wincker P., Levy M.,
RA Weissbach J., Boucher C.A.;
RT "Genome sequence of the plant pathogen Ralstonia solanacearum";
RL Nature 415:497-502(2002).
DR EMBL; AL646078; CAD17463.1;
DR InterPro; IPR001036; Acriflavin_res.
DR InterPro; IPR000731; HMGCR/patch_5TM.
DR Pfam; PF00873; ACR_tran; 1.
DR PRINTS; PR00702; ACRIFLAVINRP.
DR TIGRFAMS; TIGR00915; 2A0602; 1.
DR PROSITE; PS50156; SSD; 1.
KW Plasmid; Complete proteome.
SQ SEQUENCE 1049 AA; 111769 MW; CB59674B670089CE CRC64;

Query Match 65.4%; Score 34; DB 16; Length 1049;
Best Local Similarity 75.0%; Pred. No. 2.1e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      4 VFXGMSYS 11
Db      317 MPAGMSYS 324

RESULT 24
Q96914 PRELIMINARY; PRT; 1499 AA.
ID Q96914;
AC Q96914;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DE putative aminophospholipid translocase (Aminophospholipid-transporting
DE ATPase).
GN ATP10C.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21225279; PubMed=11326269;
RA Meguro M., Kashiwagi A., Mitsuya K., Nakao M., Kondo I., Saitoh S.,
RA Oshimura M.;
RT "A novel maternally expressed gene, ATP10C, encodes a putative
RT aminophospholipid translocase associated with Angelman syndrome.";
```

```

RL Nat. Genet. 28:19-20(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21313119; PubMed=11353404;
RA Herzig L.B.K., Kim S.J., Cook E.H. Jr., Ledbetter D.H.;
RT "The human aminophospholipid-transporting ATPase gene ATP10C maps
RT adjacent to UBE3A and exhibits similar imprinted expression.";
RL Am. J. Hum. Genet. 68:1501-1505(2001).
DR EMBL; AB051358; BAB47392.1;
DR EMBL; AY029504; AAK33100.1;
DR EMBL; AY029487; AAK33100.1; JOINED.
DR EMBL; AY029488; AAK33100.1; JOINED.
DR EMBL; AY029489; AAK33100.1; JOINED.
DR EMBL; AY029490; AAK33100.1; JOINED.
DR EMBL; AY029491; AAK33100.1; JOINED.
DR EMBL; AY029492; AAK33100.1; JOINED.
DR EMBL; AY029493; AAK33100.1; JOINED.
DR EMBL; AY029494; AAK33100.1; JOINED.
DR EMBL; AY029495; AAK33100.1; JOINED.
DR EMBL; AY029496; AAK33100.1; JOINED.
DR EMBL; AY029497; AAK33100.1; JOINED.
DR EMBL; AY029498; AAK33100.1; JOINED.
DR EMBL; AY029499; AAK33100.1; JOINED.
DR EMBL; AY029500; AAK33100.1; JOINED.
DR EMBL; AY029501; AAK33100.1; JOINED.
DR EMBL; AY029502; AAK33100.1; JOINED.
DR EMBL; AY029503; AAK33100.1; JOINED.
DR InterPro; IPR001757; ATPase_E1-E2.
DR InterPro; IPR001064; Crystallin.
DR InterPro; IPR001454; Hlgnaase/hydrlase.
DR Pfam; PF00702; Hydrolase; 1.
DR PROSITE; PS00154; ATPASE_E1_E2; UNKNOWN.1.
DR PROSITE; PS00225; CRYSTALLIN_BETAGAMMA; UNKNOWN.1.
SQ SEQUENCE 1499 AA; 167687 MW; D4996A4D0635A6BD CRC64;

Query Match 65.4%; Score 34; DB 4; Length 1499;
Best Local Similarity 72.7%; Pred. No. 3e+02;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 EEVVPXGMSYS 11
Db      469 EEVPRGGSVS 479

RESULT 25
Q8TX62 PRELIMINARY; PRT; 143 AA.
ID Q8TX62
AC Q8TX62;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Uncharacterized conserved protein.
GN MK0814.
OS Methanopyrus kandleri.
OC Archaea; Euryarchaeota; Methanopyri; Methanopyrales; Methanopyraceae;
OC Methanopyrus.
OX NCBI_TaxID=2320;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-AV19 / DSM 6324 / JCM 9639;
RX MEDLINE=21927647; PubMed=11930014;
RA Slesarev A.I., Mezhevaya K.V., Makarova K.S., Polushin N.N.,
RA Shcherbinina O.V., Shakhova V.V., Belova G.I., Aravind L.,
RA Natile D.A., Rogozin I.B., Tatusov R.L., Wolf Y.I., Stetter K.O.,
RA Malykh A.G., Koonin E.V., Kozyavkin S.A.;
RT "The complete genome of hyperthermophile Methanopyrus kandleri AV19
RT and monophyly of archaeal methanogens."
RL Proc. Natl. Acad. Sci. U.S.A. 99:4644-4649(2002).
DR EMBL; AE010372; AAM02027.1;
KW Complete proteome.
SQ SEQUENCE 143 AA; 15734 MW; 4C8B28A1FBEDD0B CRC64;

Query Match 63.5%; Score 33; DB 17; Length 143;
```

Wed Jun 11 15:45:14 2003

us-09-909-164-9.rspt

Best Local Similarity 60.0%; Pred. No. 40;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10  
||:| | |  
Db 75 EELVPOGAGY 84

Search completed: June 10, 2003, 13:46:32  
Job time : 25.7857 secs





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2003, 13:24:45 ; Search time 31.3571 seconds  
(without alignments)  
46.744 Million cell updates/sec

Title: US-09-909-164-10  
Perfect score: 52  
Sequence: 1 EENVXPXGMSYS 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues  
Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 75 summaries

Database :			
A_Geneseq_101002:*			
1:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*		
2:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*		
3:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*		
4:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*		
5:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*		
6:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*		
7:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*		
8:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*		
9:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*		
10:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*		
11:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*		
12:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*		
13:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*		
14:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*		
15:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*		
16:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*		
17:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*		
18:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*		
19:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*		
20:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*		
21:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*		
22:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*		
23:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*		

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	50	96.2	11	23	Hepatitis C virus
2	50	96.2	11	23	Hepatitis C virus
3	50	96.2	11	23	Hepatitis C virus
4	50	96.2	11	23	Hepatitis C virus
5	50	96.2	11	23	Hepatitis C virus
6	50	96.2	11	23	Hepatitis C virus
7	50	96.2	11	23	Hepatitis C virus
8	50	96.2	11	23	Hepatitis C virus
9	50	96.2	11	23	Hepatitis C virus
10	50	96.2	11	23	Hepatitis C virus

11	50	96.2	11	23	AB80568	Hepatitis C virus
12	46	88.5	11	23	AB80524	Hepatitis C virus
13	46	88.5	11	23	AB80528	Hepatitis C virus
14	46	88.5	11	23	AB80529	Hepatitis C virus
15	46	88.5	11	23	AB80561	Hepatitis C virus
16	46	88.5	11	23	AB80562	Hepatitis C virus
17	45	86.5	11	23	AB80523	Hepatitis C virus
18	45	86.5	11	23	AB80527	Hepatitis C virus
19	45	86.5	11	23	AB80535	Hepatitis C virus
20	45	86.5	11	23	AB80536	Hepatitis C virus
21	45	86.5	11	23	AB80539	Hepatitis C virus
22	45	86.5	11	23	AB80540	Hepatitis C virus
23	45	86.5	11	23	AB80558	Hepatitis C virus
24	45	86.5	11	23	AB80560	Hepatitis C virus
25	44	84.6	11	23	AB80544	Hepatitis C virus
26	44	84.6	11	23	AB80545	Hepatitis C virus
27	44	84.6	11	23	AB80549	Hepatitis C virus
28	44	84.6	11	23	AB80552	Hepatitis C virus
29	44	84.6	11	23	AB80553	Hepatitis C virus
30	42	80.8	11	23	AB80530	Hepatitis C virus
31	41	78.8	11	23	AB80538	Hepatitis C virus
32	41	78.8	11	23	AB80542	Hepatitis C virus
33	41	78.8	11	23	AB80543	Hepatitis C virus
34	40	76.9	11	23	AB80537	Hepatitis C virus
35	40	76.9	11	23	AB80541	Hepatitis C virus
36	40	76.9	11	23	AB80547	Hepatitis C virus
37	40	76.9	11	23	AB80548	Hepatitis C virus
38	40	76.9	11	23	AB80551	Hepatitis C virus
39	40	76.9	11	23	AB80556	Hepatitis C virus
40	40	76.9	11	23	AB80557	Hepatitis C virus
41	40	76.9	20	20	AAU76810	Novel human diago
42	40	76.9	1022	22	ABG03621	Novel human diago
43	40	76.9	1022	22	ABG05826	Novel human diago
44	40	76.9	1022	22	ABG08173	Hepatitis C virus
45	39	75.0	11	23	AB80546	Hepatitis C virus
46	39	75.0	11	23	AB80550	Hepatitis C virus
47	39	75.0	11	23	AB80554	Hepatitis C virus
48	39	75.0	11	23	AB80555	Hepatitis C virus
49	38	73.1	11	23	AB80533	Hepatitis C virus
50	38	73.1	11	23	AB80534	Hepatitis C virus
51	38	73.1	3472	21	AAV90913	Cenarchaeum symbio
52	37	71.2	11	23	AB80531	Hepatitis C virus
53	37	71.2	11	23	AB80532	Hepatitis C virus
54	36	69.2	244	21	AA12881	Murine JNK3 bindin
55	36	69.2	484	21	AA12882	Murine JNK3 bindin
56	35	67.3	11	18	AAW99288	Peptide N424 from
57	34	65.4	842	21	AAV44359	P. chrysogenum sut
58	34	65.4	947	21	AAV45105	Pinus radiata cell
59	34	65.4	1070	22	AAU14378	Human novel protei
60	33	63.5	12	21	AAV83772	HCV NS3 protease s
61	33	63.5	12	21	AAV83774	HCV NS3 protease s
62	33	63.5	13	18	AAW99276	Peptide D4 from WO
63	33	63.5	13	18	AAW99286	Peptide 5 used in
64	33	63.5	14	18	AAW99275	Peptide D3 from WO
65	33	63.5	14	18	AAW99277	Peptide C0 from WO
66	33	63.5	14	18	AAW33285	Peptide 4 used in
67	33	63.5	14	18	AAW33287	Peptide 6 used in
68	33	63.5	16	18	AAW99274	Peptide D2 from WO
69	33	63.5	16	18	AAW33284	Peptide 3 used in
70	33	63.5	18	18	AAW99273	Peptide D1 from WO
71	33	63.5	18	18	AAW33283	Peptide 2 used in
72	33	63.5	18	19	AAW71281	Cleavable substrate
73	33	63.5	20	18	AAW99272	Peptide PS from WO
74	33	63.5	20	18	AAW33282	Peptide 1 used in
75	33	63.5	20	20	AAV14511	HCV NS3 protease t

ALIGNMENTS

RESULT 1  
AB80521  
ID AB80521 standard; peptide; 11 AA.





XX PA (CORV-) CORVAS INT INC.  
 XX PI Lim-wilby M, Levy OE, Brunck TK;  
 XX WPI; 2002-361643/39.  
 XX Novel peptide compound having hepatitis C virus protease inhibitory  
 PT activity useful for treating disorders associated with hepatitis C  
 PT virus protease -  
 XX  
 PS Claim 17; Page 65; 69pp; English.  
 XX  
 CC The sequence represents a peptide compound of the invention having  
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
 CC invention are alpha-ketoamide peptide analogues. The peptides have  
 CC virucide activity, and are useful for treating and in the manufacture of  
 CC a medicament to treat disorders associated with HCV protease. A  
 CC pharmaceutical composition comprising the peptide as an active ingredient  
 CC is useful for treating disorders associated with hepatitis C virus.  
 XX  
 SQ Sequence 11 AA;

Query Match 96.2%; Score 50; DB 23; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 0.0011;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 EEVVPXGMSYS 11  
 |||||  
 Db 1 EEVVPXGMSYS 11

RESULT 6  
 ABB80563  
 ID ABB80563 standard; peptide; 11 AA.  
 XX AC ABB80563;  
 XX  
 DT 08-OCT-2002 (first entry)  
 DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #43.  
 XX  
 KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
 KW virucide.  
 XX  
 OS Synthetic.  
 FH Key Location/Qualifiers  
 FT Modified-site 1 /note= "N-terminal acetyl"  
 FT Modified-site 6  
 FT Modified-site 6 /note= "Valyl carbonyl forming keto-amide linkage with  
 FT residue 7"  
 FT Modified-site 11 /note= "C-terminal amide"  
 FT  
 PN WO200208251-A2.  
 XX  
 PD 31-JAN-2002.  
 XX  
 PF 19-JUL-2001; 2001WO-US23169.  
 XX  
 PR 21-JUL-2000; 2000US-220101P.  
 XX  
 PA (CORV-) CORVAS INT INC.  
 XX  
 PI Lim-wilby M, Levy OE, Brunck TK;  
 XX WPI; 2002-361643/39.  
 XX Novel peptide compound having hepatitis C virus protease inhibitory  
 PT activity useful for treating disorders associated with hepatitis C  
 PT virus protease -

XX PS Claim 17; Page 65; 69pp; English.  
 XX

CC The sequence represents a peptide compound of the invention having  
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
 CC invention are alpha-ketoamide peptide analogues. The peptides have  
 CC virucide activity, and are useful for treating and in the manufacture of  
 CC a medicament to treat disorders associated with HCV protease. A  
 CC pharmaceutical composition comprising the peptide as an active ingredient  
 CC is useful for treating disorders associated with hepatitis C virus.  
 XX

SQ Sequence 11 AA;

Query Match 96.2%; Score 50; DB 23; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 0.0011;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 EEVVPXGMSYS 11  
 |||||  
 Db 1 EEVVPXGMSYS 11

RESULT 7  
 ABB80564  
 ID ABB80564 standard; peptide; 11 AA.  
 XX AC ABB80564;  
 XX

DT 08-OCT-2002 (first entry)

DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #44.

KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
 KW virucide.

XX OS Synthetic.

FH Key Location/Qualifiers  
 FT Modified-site 1 /note= "N-terminal acetyl"  
 FT Modified-site 6  
 FT Modified-site 6 /note= "Leucyl carbonyl forming keto-amide linkage with  
 FT residue 7"  
 FT Modified-site 11 /note= "C-terminal amide"  
 FT

PN WO200208251-A2.

XX 31-JAN-2002.

PF 19-JUL-2001; 2001WO-US23169.

PR 21-JUL-2000; 2000US-220101P.

XX (CORV-) CORVAS INT INC.

PI Lim-wilby M, Levy OE, Brunck TK;

XX WPI; 2002-361643/39.

XX Novel peptide compound having hepatitis C virus protease inhibitory  
 PT activity useful for treating disorders associated with hepatitis C  
 PT virus protease -

XX PS Claim 17; Page 65; 69pp; English.  
 XX

CC The sequence represents a peptide compound of the invention having  
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
 CC invention are alpha-ketoamide peptide analogues. The peptides have  
 CC virucide activity, and are useful for treating and in the manufacture of  
 CC a medicament to treat disorders associated with HCV protease. A  
 CC pharmaceutical composition comprising the peptide as an active ingredient  
 CC is useful for treating disorders associated with hepatitis C virus.

[illegible]



us-09-909-164-10.rag

Wed Jun 11 15:41:48 2003

PI Lim-wilby M, Levy OE, Brunck TK;  
 XX WPI; 2002-361643/39.  
 XX Novel peptide compound having hepatitis C virus protease inhibitory  
 PT activity useful for treating disorders associated with hepatitis C  
 PT virus protease -  
 XX  
 XX Claim 17; Page 64; 69pp; English.  
 PS  
 XX The sequence represents a peptide compound of the invention having  
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
 CC invention are alpha-ketoamide peptide analogues. The peptides have  
 CC virucide activity, and are useful for treating and in the manufacture of  
 CC a medicament to treat disorders associated with HCV protease. A  
 CC pharmaceutical composition comprising the peptide as an active ingredient  
 CC is useful for treating disorders associated with hepatitis C virus.  
 XX  
 XX Sequence 11 AA;  
 SQ  
 Query Match 88.5%; Score 46; DB 23; Length 11;  
 Best Local Similarity 90.9%; Pred. No. 0.0075;  
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 EEVVPXGMSYS 11  
 |||||  
 DB 1 EEVVPXGMDYS 11  
 |||||  
 RESULT 14  
 ABB80529  
 ID ABB80529 standard; peptide; 11 AA.  
 XX  
 AC ABB80529;  
 XX  
 DT 08-OCT-2002 (first entry)  
 XX  
 DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #9.  
 XX  
 KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
 KW virucide.  
 XX  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Modified-site 1 /note= "N-terminal acetyl"  
 FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with  
 FT residue 7"  
 FT Misc-difference 8 /note= "D-form residue"  
 FT Misc-difference 9 /note= "D-form residue"  
 FT Modified-site 11 /note= "C-terminal amide"  
 FT  
 XX WO200208251-A2.  
 XX  
 XX 31-JAN-2002.  
 XX  
 XX 19-JUL-2001; 2001WO-US23169.  
 XX  
 XX 21-JUL-2000; 2000US-220101P.  
 XX  
 XX (CORV-) CORVAS INT INC.  
 XX  
 XX Lim-wilby M, Levy OE, Brunck TK;  
 XX  
 XX WPI; 2002-361643/39.  
 XX Novel peptide compound having hepatitis C virus protease inhibitory  
 PT activity useful for treating disorders associated with hepatitis C.

PD 31-JAN-2002.  
 XX  
 PF 19-JUL-2001; 2001WO-US23169.  
 XX  
 XX 21-JUL-2000; 2000US-220101P.  
 PR  
 XX (CORV-) CORVAS INT INC.  
 XX  
 XX Lim-wilby M, Levy OE, Brunck TK;  
 XX  
 XX WPI; 2002-361643/39.  
 XX Novel peptide compound having hepatitis C virus protease inhibitory  
 PT activity useful for treating disorders associated with hepatitis C  
 PT virus protease -  
 XX  
 XX Claim 17; Page 64; 69pp; English.  
 PS  
 XX The sequence represents a peptide compound of the invention having  
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
 CC invention are alpha-ketoamide peptide analogues. The peptides have  
 CC virucide activity, and are useful for treating and in the manufacture of  
 CC a medicament to treat disorders associated with HCV protease. A  
 CC pharmaceutical composition comprising the peptide as an active ingredient  
 CC is useful for treating disorders associated with hepatitis C virus.  
 XX  
 XX Sequence 11 AA;  
 SQ  
 Query Match 88.5%; Score 46; DB 23; Length 11;  
 Best Local Similarity 90.9%; Pred. No. 0.0075;  
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 EEVVPXGMSYS 11  
 |||||  
 DB 1 EEVVPXGMDYS 11  
 |||||  
 RESULT 13  
 ABB80528  
 ID ABB80528 standard; peptide; 11 AA.  
 XX  
 AC ABB80528;  
 XX  
 DT 08-OCT-2002 (first entry)  
 XX  
 DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #8.  
 XX  
 KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
 KW virucide.  
 XX  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Modified-site 1 /note= "N-terminal acetyl"  
 FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with  
 FT residue 7"  
 FT Misc-difference 8 /note= "D-form residue"  
 FT Modified-site 11 /note= "C-terminal amide"  
 FT  
 XX WO200208251-A2.  
 XX  
 XX 31-JAN-2002.  
 XX  
 XX 19-JUL-2001; 2001WO-US23169.  
 XX  
 XX 21-JUL-2000; 2000US-220101P.  
 XX  
 XX (CORV-) CORVAS INT INC.

PT virus protease -  
 XX Claim 17; Page 64; 69pp; English.  
 XX  
 CC The sequence represents a peptide compound of the invention having  
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
 CC invention are alpha-ketoamide peptide analogues. The peptides have  
 CC virucide activity, and are useful for treating and in the manufacture of  
 CC a medicament to treat disorders associated with HCV protease. A  
 CC pharmaceutical composition comprising the peptide as an active ingredient  
 CC is useful for treating disorders associated with hepatitis C virus.  
 XX  
 SQ Sequence 11 AA;  
 Query Match 88.5%; Score 46; DB 23; Length 11;  
 Best Local Similarity 90.9%; Pred. No. 0.0075;  
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 EEVVPXGMSYS 11  
 Db 1 EEVVPXGMDYS 11  
 ID ABB80561 standard; peptide; 11 AA.  
 AC ABB80561;  
 XX 08-OCT-2002 (first entry)  
 DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #41.  
 XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
 KW virucide.  
 KW Synthetic.  
 OS  
 XX Key Location/Qualifiers  
 FT Modified-site 1 /note= "N-terminal acetyl"  
 FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with  
 FT Misc-difference 8 residue 7"  
 FT Modified-site 8 /note= "D-form residue"  
 FT Modified-site 8 /note= "Oxymethionine"  
 FT Misc-difference 9 /note= "D-form residue"  
 FT Modified-site 11 /note= "C-terminal amide"  
 FT Modified-site 11 /note= "C-terminal amide"  
 XX WO200208251-A2.  
 PN 31-JAN-2002.  
 XX 19-JUL-2001; 2001WO-US23169.  
 XX 21-JUL-2000; 2000US-220101P.  
 XX (CORV-) CORVAS INT INC.  
 XX Lim-wilby M, Levy OE, Brunck TK;  
 PI WPI; 2002-361643/39.  
 DR Novel peptide compound having hepatitis C virus protease inhibitory  
 PT activity useful for treating disorders associated with hepatitis C  
 PT virus protease -  
 XX  
 PS Claim 17; Page 65; 69pp; English.  
 CC The sequence represents a peptide compound of the invention having  
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
 CC invention are alpha-ketoamide peptide analogues. The peptides have  
 CC virucide activity, and are useful for treating and in the manufacture of  
 CC a medicament to treat disorders associated with HCV protease. A  
 CC pharmaceutical composition comprising the peptide as an active ingredient  
 CC is useful for treating disorders associated with hepatitis C virus.

CC invention are alpha-ketoamide peptide analogues. The peptides have  
 CC virucide activity, and are useful for treating and in the manufacture of  
 CC a medicament to treat disorders associated with HCV protease. A  
 CC pharmaceutical composition comprising the peptide as an active ingredient  
 CC is useful for treating disorders associated with hepatitis C virus.  
 XX  
 SQ Sequence 11 AA;  
 Query Match 88.5%; Score 46; DB 23; Length 11;  
 Best Local Similarity 90.9%; Pred. No. 0.0075;  
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 EEVVPXGMSYS 11  
 Db 1 EEVVPXGMDYS 11  
 ID ABB80562 standard; peptide; 11 AA.  
 AC ABB80562;  
 XX 08-OCT-2002 (first entry)  
 DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #42.  
 XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
 KW virucide.  
 KW Synthetic.  
 OS  
 XX Key Location/Qualifiers  
 FT Modified-site 1 /note= "N-terminal acetyl"  
 FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with  
 FT Misc-difference 8 residue 7"  
 FT Modified-site 8 /note= "D-form residue"  
 FT Modified-site 8 /note= "Oxymethionine"  
 FT Misc-difference 9 /note= "D-form residue"  
 FT Modified-site 11 /note= "C-terminal amide"  
 FT Modified-site 11 /note= "C-terminal amide"  
 XX WO200208251-A2.  
 PN 31-JAN-2002.  
 XX 19-JUL-2001; 2001WO-US23169.  
 XX 21-JUL-2000; 2000US-220101P.  
 XX (CORV-) CORVAS INT INC.  
 XX Lim-wilby M, Levy OE, Brunck TK;  
 PI WPI; 2002-361643/39.  
 DR Novel peptide compound having hepatitis C virus protease inhibitory  
 PT activity useful for treating disorders associated with hepatitis C  
 PT virus protease -  
 XX  
 PS Claim 17; Page 65; 69pp; English.  
 CC The sequence represents a peptide compound of the invention having  
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
 CC invention are alpha-ketoamide peptide analogues. The peptides have  
 CC virucide activity, and are useful for treating and in the manufacture of  
 CC a medicament to treat disorders associated with HCV protease. A  
 CC pharmaceutical composition comprising the peptide as an active ingredient  
 CC is useful for treating disorders associated with hepatitis C virus.



us-09-909-164-10.rag

Wed Jun 11 15:41:48 2003

CC is useful for treating disorders associated with hepatitis C virus.  
 XX Sequence 11 AA;

Query Match 88.5%; Score 46; DB 23; Length 11;  
 Best Local Similarity 90.9%; Pred. No. 0.0075; 1; Indels 0; Gaps 0;  
 Matches 10; Conservative 0; Mismatches 1;

QY 1 EEVVPXGMSYS 11  
 DB 1 EEVVPXGMDYS 11

RESULT 17  
 ABB80523  
 ID ABB80523 standard; peptide; 11 AA.  
 XX ABB80523;  
 XX 08-OCT-2002 (first entry)  
 XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #3.  
 XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
 XX virucide.  
 XX Synthetic.  
 XX Key Location/Qualifiers  
 FT Modified-site 1 /note= "N-terminal acetyl"  
 FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with  
 FT residue 7"  
 FT Misc-difference 9 /note= "D-form residue"  
 FT Modified-site 11 /note= "C-terminal amide"  
 FT WO200208251-A2.  
 XX 31-JAN-2002.  
 XX 19-JUL-2001; 2001WO-US23169.  
 XX 21-JUL-2000; 2000US-220101P.  
 XX (CORV-) CORVAS INT INC.  
 XX Lim-wilby M, Levy OE, Brunck TK;  
 XX WPI; 2002-361643/39.  
 XX Novel peptide compound having hepatitis C virus protease inhibitory  
 XX activity useful for treating disorders associated with hepatitis C  
 XX virus protease  
 XX Claim 17; Page 64; 69pp; English.

The sequence represents a peptide compound of the invention having  
 hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
 invention are alpha-ketoamide peptide analogues. The peptides have  
 virucide activity, and are useful for treating and in the manufacture of  
 a medicament to treat disorders associated with HCV protease. A  
 pharmaceutical composition comprising the peptide as an active ingredient  
 is useful for treating disorders associated with hepatitis C virus.

Query Match 86.5%; Score 45; DB 23; Length 11;  
 Best Local Similarity 90.9%; Pred. No. 0.012;  
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 Sequence 11 AA;

QY 1 EEVVPXGMSYS 11  
 DB 1 EEVVPXGMDYS 11

RESULT 19  
 ABB80535  
 ID ABB80535 standard; peptide; 11 AA.

QY 1 EEVVPXGMSYS 11  
 DB 1 EEVVPXGMHYS 11

RESULT 18  
 ABB80527  
 ID ABB80527 standard; peptide; 11 AA.

XX ABB80527;  
 XX 08-OCT-2002 (first entry)  
 XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #7.  
 XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
 XX virucide.  
 XX Synthetic.  
 XX Key Location/Qualifiers  
 FT Modified-site 1 /note= "N-terminal acetyl"  
 FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with  
 FT residue 7"  
 FT Misc-difference 8 /note= "D-form residue"  
 FT Modified-site 11 /note= "C-terminal amide"  
 FT WO200208251-A2.  
 XX 31-JAN-2002.  
 XX 19-JUL-2001; 2001WO-US23169.  
 XX 21-JUL-2000; 2000US-220101P.  
 XX (CORV-) CORVAS INT INC.  
 XX Lim-wilby M, Levy OE, Brunck TK;  
 XX WPI; 2002-361643/39.  
 XX Novel peptide compound having hepatitis C virus protease inhibitory  
 XX activity useful for treating disorders associated with hepatitis C  
 XX virus protease  
 XX Claim 17; Page 64; 69pp; English.

The sequence represents a peptide compound of the invention having  
 hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
 invention are alpha-ketoamide peptide analogues. The peptides have  
 virucide activity, and are useful for treating and in the manufacture of  
 a medicament to treat disorders associated with HCV protease. A  
 pharmaceutical composition comprising the peptide as an active ingredient  
 is useful for treating disorders associated with hepatitis C virus.

Query Match 86.5%; Score 45; DB 23; Length 11;  
 Best Local Similarity 90.9%; Pred. No. 0.012;  
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 Sequence 11 AA;

QY 1 EEVVPXGMSYS 11  
 DB 1 EEVVPXGMHYS 11

RESULT 19  
 ABB80535  
 ID ABB80535 standard; peptide; 11 AA.

XX AC ABB80535;  
 XX DT 08-OCT-2002 (first entry)  
 XX DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #15.  
 XX KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
 XX KW virucide.  
 XX OS Synthetic.  
 XX FH Key Location/Qualifiers  
 XX FT Modified-site 1 /note= "N-terminal acetyl"  
 XX FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with  
 XX FT residue 7"  
 XX OS Synthetic.  
 XX FH Key Location/Qualifiers  
 XX FT Modified-site 1 /note= "N-terminal acetyl"  
 XX FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with  
 XX FT residue 7"  
 XX FT Modified-site 11 /note= "C-terminal amide"  
 XX PN WO200208251-A2.  
 XX PD 31-JAN-2002.  
 XX PF 19-JUL-2001; 2001WO-US23169.  
 XX PR 21-JUL-2000; 2000US-220101P.  
 XX PA (CORV-) CORVAS INT INC.  
 XX PI Lim-wilby M, Levy OE, Brunck TK;  
 XX DR WPI; 2002-361643/39.  
 XX CC Novel peptide compound having hepatitis C virus protease inhibitory  
 XX PT activity useful for treating disorders associated with hepatitis C  
 XX PT virus protease -  
 XX PS Claim 17; Page 64; 69pp; English.  
 XX CC The sequence represents a peptide compound of the invention having  
 XX CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
 XX CC invention are alpha-ketoamide peptide analogues. The peptides have  
 XX CC virucide activity, and are useful for treating and in the manufacture of  
 XX CC a medicament to treat disorders associated with HCV protease. A  
 XX CC pharmaceutical composition comprising the peptide as an active ingredient  
 XX CC is useful for treating disorders associated with hepatitis C virus.  
 XX SQ Sequence 11 AA;  
 XX Query Match 86.5%; Score 45; DB 23; Length 11;  
 XX Best Local Similarity 90.9%; Pred. No. 0.012;  
 XX Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 XX QY 1 EEVVPXGMSYS 11  
 XX DB ||||| |||  
 XX 1 EEVVPXGMSYS 11  
 XX RESULT 20  
 XX ID ABB80536 standard; peptide; 11 AA.  
 XX AC ABB80536;  
 XX XX 08-OCT-2002 (first entry)  
 XX DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #16.  
 XX KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
 XX KW virucide.

OS Synthetic.  
 XX Key Location/Qualifiers  
 XX FT Modified-site 1 /note= "N-terminal acetyl"  
 XX FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with  
 XX FT residue 7"  
 XX FT Misc-difference 9 /note= "D-form residue"  
 XX FT Modified-site 11 /note= "C-terminal amide"  
 XX FT WO200208251-A2.  
 XX PN 31-JAN-2002.  
 XX PD 19-JUL-2001; 2001WO-US23169.  
 XX PF 21-JUL-2000; 2000US-220101P.  
 XX PR (CORV-) CORVAS INT INC.  
 XX PA Lim-wilby M, Levy OE, Brunck TK;  
 XX XX WPI; 2002-361643/39.  
 XX CC Novel peptide compound having hepatitis C virus protease inhibitory  
 XX PT activity useful for treating disorders associated with hepatitis C  
 XX PT virus protease -  
 XX PS Claim 17; Page 64; 69pp; English.  
 XX CC The sequence represents a peptide compound of the invention having  
 XX CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
 XX CC invention are alpha-ketoamide peptide analogues. The peptides have  
 XX CC virucide activity, and are useful for treating and in the manufacture of  
 XX CC a medicament to treat disorders associated with HCV protease. A  
 XX CC pharmaceutical composition comprising the peptide as an active ingredient  
 XX CC is useful for treating disorders associated with hepatitis C virus.  
 XX SQ Sequence 11 AA;  
 XX Query Match 86.5%; Score 45; DB 23; Length 11;  
 XX Best Local Similarity 90.9%; Pred. No. 0.012;  
 XX Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 XX QY 1 EEVVPXGMSYS 11  
 XX DB ||||| |||  
 XX 1 EEVVPXGMSYS 11  
 XX RESULT 21  
 XX ID ABB80539 standard; peptide; 11 AA.  
 XX AC ABB80539;  
 XX XX 08-OCT-2002 (first entry)  
 XX DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #19.  
 XX KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
 XX KW virucide.  
 XX OS Synthetic.  
 XX FH Key Location/Qualifiers  
 XX FT Modified-site 1 /note= "N-terminal acetyl"  
 XX FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with  
 XX FT residue 7"

XX 31-JAN-2002.  
 XX PD  
 XX XX  
 XX PF 19-JUL-2001; 2001WO-US23169.  
 XX FT  
 XX PR 21-JUL-2000; 2000US-220101P.  
 XX PA (CORV-) CORVAS INT INC.  
 XX PI LIm-wilby M, Levy OE, Brunc TK;  
 XX XX  
 XX XX WPI; 2002-361643/39.  
 XX DR  
 XX PT Novel peptide compound having hepatitis C virus protease inhibitory  
 XX PT activity useful for treating disorders associated with hepatitis C  
 XX PT virus protease -  
 XX XX  
 XX PS Claim 17; Page 65; 69pp; English.  
 XX CC  
 XX CC The sequence represents a peptide compound of the invention having  
 XX CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
 XX CC invention are alpha-ketoamide peptide analogues. The peptides have  
 XX CC virucide activity, and are useful for treating and in the manufacture of  
 XX CC a medicament to treat disorders associated with HCV protease. A  
 XX CC pharmaceutical composition comprising the peptide as an active ingredient  
 XX CC is useful for treating disorders associated with hepatitis C virus.  
 XX SQ Sequence 11 AA;  
 Query Match 86.5%; Score 45; DB 23; Length 11;  
 Best Local Similarity 90.9%; Pred. No. 0.012; 1; Indels 0; Gaps 0;  
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 EEVVPXGMSYS 11  
 DB 1 EEVVPXGQSYS 11  
 RESULT 23  
 ID ABB80558  
 XX ID ABB80558 standard; peptide; 11 AA.  
 XX AC ABB80558;  
 XX DT 08-OCT-2002 (first entry)  
 XX DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #38.  
 XX KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
 XX KW virucide.  
 XX OS Synthetic.  
 XX PH Key Location/Qualifiers  
 FT Modified-site 1 /note= "N-terminal acetyl"  
 FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with  
 FT Modified-site 8 residue 7"  
 FT Modified-site 11 /note= "Oxymethionine"  
 FT Modified-site 11 /note= "C-terminal amide"  
 XX WO200208251-A2.  
 XX 31-JAN-2002.  
 XX PD  
 XX PF 19-JUL-2001; 2001WO-US23169.  
 XX PR 21-JUL-2000; 2000US-220101P.  
 XX PA (CORV-) CORVAS INT INC.

FT Misc-difference 8 /note= "D-form residue"  
 FT Modified-site 11 /note= "C-terminal amide"  
 XX WO200208251-A2.  
 XX PN  
 XX PD 31-JAN-2002.  
 XX XX  
 XX PF 19-JUL-2001; 2001WO-US23169.  
 XX PR 21-JUL-2000; 2000US-220101P.  
 XX PA (CORV-) CORVAS INT INC.  
 XX PI LIm-wilby M, Levy OE, Brunc TK;  
 XX XX  
 XX XX WPI; 2002-361643/39.  
 XX DR  
 XX PT Novel peptide compound having hepatitis C virus protease inhibitory  
 XX PT activity useful for treating disorders associated with hepatitis C  
 XX PT virus protease -  
 XX XX  
 XX PS Claim 17; Page 65; 69pp; English.  
 XX CC  
 XX CC The sequence represents a peptide compound of the invention having  
 XX CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
 XX CC invention are alpha-ketoamide peptide analogues. The peptides have  
 XX CC virucide activity, and are useful for treating and in the manufacture of  
 XX CC a medicament to treat disorders associated with HCV protease. A  
 XX CC pharmaceutical composition comprising the peptide as an active ingredient  
 XX CC is useful for treating disorders associated with hepatitis C virus.  
 XX SQ Sequence 11 AA;  
 Query Match 86.5%; Score 45; DB 23; Length 11;  
 Best Local Similarity 90.9%; Pred. No. 0.012; 1; Indels 0; Gaps 0;  
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 EEVVPXGMSYS 11  
 DB 1 EEVVPXGQSYS 11  
 RESULT 22  
 ID ABB80540  
 XX ID ABB80540 standard; peptide; 11 AA.  
 XX AC ABB80540;  
 XX DT 08-OCT-2002 (first entry)  
 XX DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #20.  
 XX KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
 XX KW virucide.  
 XX OS Synthetic.  
 XX PH Key Location/Qualifiers  
 FT Modified-site 1 /note= "N-terminal acetyl"  
 FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with  
 FT Modified-site 8 residue 7"  
 FT Misc-difference 8 /note= "D-form residue"  
 FT Misc-difference 9 /note= "D-form residue"  
 FT Modified-site 11 /note= "C-terminal amide"  
 XX WO200208251-A2.  
 XX PN

XX Llm-wilby M, Levy OE, Brunck TK;  
XX WPI; 2002-361643/39.  
XX Novel peptide compound having hepatitis C virus protease inhibitory  
XX activity useful for treating disorders associated with hepatitis C  
XX virus protease -  
XX Claim 17; Page 65; 69pp; English.  
XX The sequence represents a peptide compound of the invention having  
XX hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
XX invention are alpha-ketoamide peptide analogues. The peptides have  
XX virucide activity, and are useful for treating and in the manufacture of  
XX a medicament to treat disorders associated with HCV protease. A  
XX pharmaceutical composition comprising the peptide as an active ingredient  
XX is useful for treating disorders associated with hepatitis C virus.  
XX Sequence 11 AA;  
SQ

Query Match 86.5%; Score 45; DB 23; Length 11;  
Best Local Similarity 90.9%; Pred. No. 0.012;  
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 EEVVPXGMSYS 11  
|||||  
Db 1 EEVVPXGMHYS 11

RESULT 24  
ABB80560  
ID ABB80560 standard; peptide; 11 AA.  
XX AC ABB80560;  
XX DT 08-OCT-2002 (first entry)  
XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #40.  
XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
XX virucide.  
XX Synthetic.  
XX Key Location/Qualifiers  
FH Modified-site 1 /note= "N-terminal acetyl"  
FT Modified-site 6  
FT Modified-site 6  
FT Modified-site 11 /note= "Norvalyl carbonyl forming keto-amide linkage with  
FT residue 7"  
FT Misc-difference 8 /note= "D-form residue"  
FT Modified-site 8 /note= "D-form residue"  
FT Misc-difference 9 /note= "Oxymethionine"  
FT Modified-site 11 /note= "D-form residue"  
FT Modified-site 11 /note= "C-terminal amide"  
XX WO200208251-A2.  
XX 31-JAN-2002.  
XX 19-JUL-2001; 2001WO-US23169.  
XX 21-JUL-2000; 2000US-220101P.  
XX (CORV-) CORVAS INT INC.  
XX Llm-wilby M, Levy OE, Brunck TK;  
XX WPI; 2002-361643/39.  
XX Novel peptide compound having hepatitis C virus protease inhibitory  
XX activity useful for treating disorders associated with hepatitis C  
XX virus protease -  
XX Claim 17; Page 65; 69pp; English.  
XX The sequence represents a peptide compound of the invention having  
XX hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
XX invention are alpha-ketoamide peptide analogues. The peptides have  
XX virucide activity, and are useful for treating and in the manufacture of  
XX a medicament to treat disorders associated with HCV protease. A  
XX pharmaceutical composition comprising the peptide as an active ingredient  
XX is useful for treating disorders associated with hepatitis C virus.  
XX Sequence 11 AA;  
SQ

Query Match 86.5%; Score 45; DB 23; Length 11;  
Best Local Similarity 90.9%; Pred. No. 0.012;  
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 EEVVPXGMSYS 11  
|||||  
Db 1 EEVVPXGMHYS 11

RESULT 25  
ABB80544  
ID ABB80544 standard; peptide; 11 AA.  
XX AC ABB80544;  
XX DT 08-OCT-2002 (first entry)  
XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #24.  
XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
XX virucide.  
XX Synthetic.  
XX Key Location/Qualifiers  
FH Modified-site 1 /note= "N-terminal acetyl"  
FT Modified-site 6  
FT Modified-site 6  
FT Modified-site 11 /note= "Norvalyl carbonyl forming keto-amide linkage with  
FT residue 7"  
FT Modified-site 11 /note= "C-terminal amide"  
XX WO200208251-A2.  
XX 31-JAN-2002.  
XX 19-JUL-2001; 2001WO-US23169.  
XX 21-JUL-2000; 2000US-220101P.  
XX (CORV-) CORVAS INT INC.  
XX Llm-wilby M, Levy OE, Brunck TK;  
XX WPI; 2002-361643/39.  
XX Novel peptide compound having hepatitis C virus protease inhibitory  
XX activity useful for treating disorders associated with hepatitis C  
XX virus protease -  
XX Claim 17; Page 65; 69pp; English.  
XX The sequence represents a peptide compound of the invention having  
XX hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
XX invention are alpha-ketoamide peptide analogues. The peptides have  
XX virucide activity, and are useful for treating and in the manufacture of  
XX a medicament to treat disorders associated with HCV protease. A  
XX pharmaceutical composition comprising the peptide as an active ingredient  
XX is useful for treating disorders associated with hepatitis C virus.  
XX Sequence 11 AA;  
SQ

XX Novel peptide compound having hepatitis C virus protease inhibitory  
XX activity useful for treating disorders associated with hepatitis C  
XX virus protease -  
XX Claim 17; Page 65; 69pp; English.  
XX The sequence represents a peptide compound of the invention having  
XX hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
XX invention are alpha-ketoamide peptide analogues. The peptides have  
XX virucide activity, and are useful for treating and in the manufacture of  
XX a medicament to treat disorders associated with HCV protease. A  
XX pharmaceutical composition comprising the peptide as an active ingredient  
XX is useful for treating disorders associated with hepatitis C virus.  
XX Sequence 11 AA;  
SQ

Query Match 86.5%; Score 45; DB 23; Length 11;  
Best Local Similarity 90.9%; Pred. No. 0.012;  
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 EEVVPXGMSYS 11  
|||||  
Db 1 EEVVPXGMHYS 11

RESULT 25  
ABB80544  
ID ABB80544 standard; peptide; 11 AA.  
XX AC ABB80544;  
XX DT 08-OCT-2002 (first entry)  
XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #24.  
XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
XX virucide.  
XX Synthetic.  
XX Key Location/Qualifiers  
FH Modified-site 1 /note= "N-terminal acetyl"  
FT Modified-site 6  
FT Modified-site 6  
FT Modified-site 11 /note= "Norvalyl carbonyl forming keto-amide linkage with  
FT residue 7"  
FT Modified-site 11 /note= "C-terminal amide"  
XX WO200208251-A2.  
XX 31-JAN-2002.  
XX 19-JUL-2001; 2001WO-US23169.  
XX 21-JUL-2000; 2000US-220101P.  
XX (CORV-) CORVAS INT INC.  
XX Llm-wilby M, Levy OE, Brunck TK;  
XX WPI; 2002-361643/39.  
XX Novel peptide compound having hepatitis C virus protease inhibitory  
XX activity useful for treating disorders associated with hepatitis C  
XX virus protease -  
XX Claim 17; Page 65; 69pp; English.  
XX The sequence represents a peptide compound of the invention having  
XX hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
XX invention are alpha-ketoamide peptide analogues. The peptides have  
XX virucide activity, and are useful for treating and in the manufacture of  
XX a medicament to treat disorders associated with HCV protease. A  
XX pharmaceutical composition comprising the peptide as an active ingredient  
XX is useful for treating disorders associated with hepatitis C virus.  
XX Sequence 11 AA;  
SQ

us-09-909-164-10.rag

Wed Jun 11 15:41:48 2003

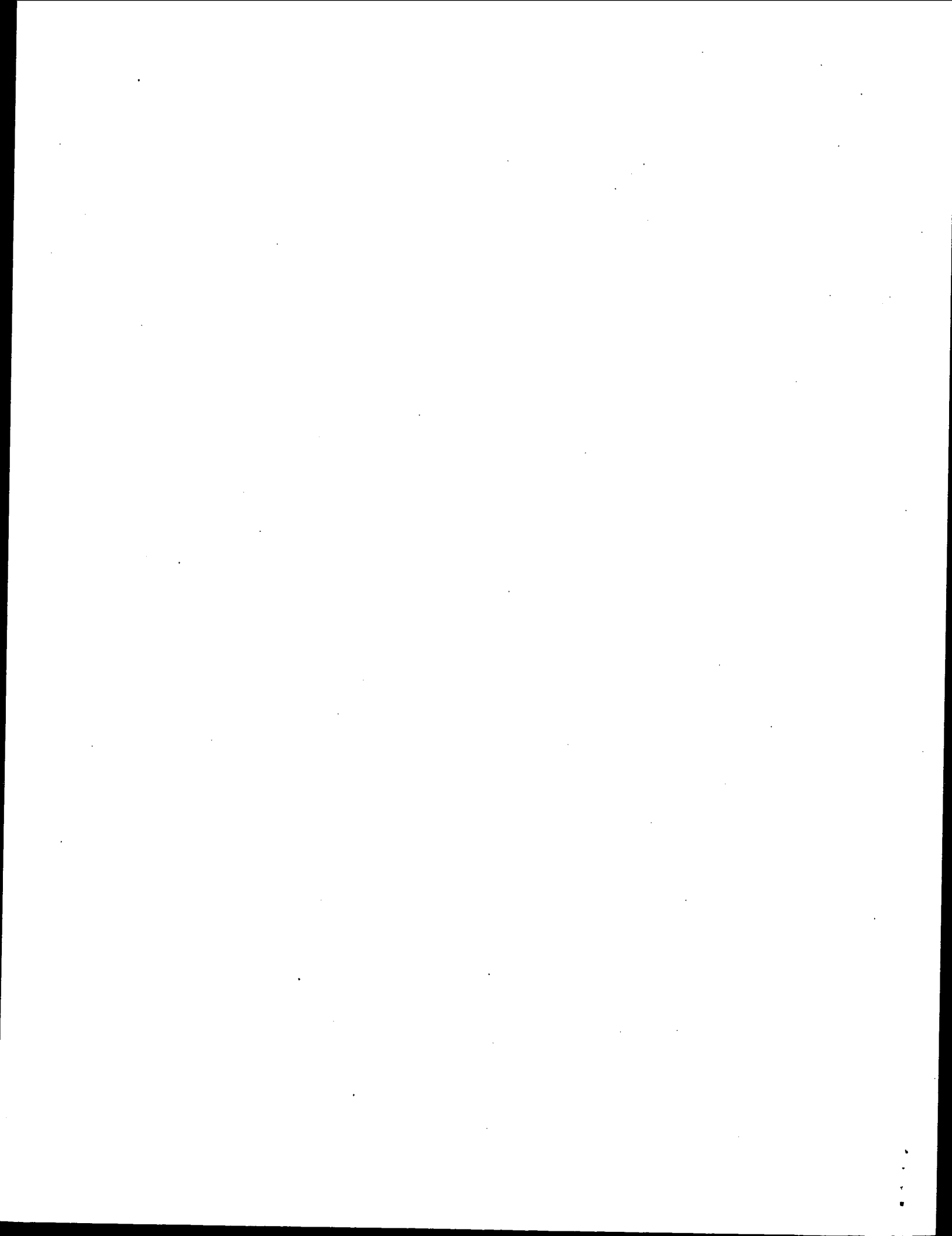
CC virucide activity, and are useful for treating and in the manufacture of  
 CC a medicament to treat disorders associated with HCV protease. A  
 CC pharmaceutical composition comprising the peptide as an active ingredient  
 CC is useful for treating disorders associated with hepatitis C virus.

XX  
 SQ Sequence 11 AA;

Query Match 84.6%; Score 44; DB 23; Length 11;  
 Best Local Similarity 90.9%; Pred. No. 0.019;  
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 EEVVPXGMSYS 11  
 |||||  
 Db 1 EEVVPXGTSYS 11

Search completed: June 10, 2003, 13:39:08  
 Job time : 31.3571 secs



Wed Jun 11 15:41:53 2003

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model  
Run on: June 10, 2003, 13:31:45 ; Search time 9.64286 Seconds  
(without alignments)  
33.564 Million cell updates/sec

Title: US-09-909-164-10  
Perfect score: 52  
Sequence: 1 BEVVPXGMSYS 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues 262574  
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 75 summaries

Database : Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PTCUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	34	65.4	947	4	US-09-228-986-73
2	33	63.5	45	2	US-08-637-759B-236
3	33	63.5	45	3	US-08-871-355A-236
4	33	63.5	45	4	US-09-201-945-236
5	33	63.5	65	6	5177197-51
6	33	63.5	410	6	5177197-1
7	33	63.5	1394	6	5177197-30
8	32	61.5	10	4	US-09-357-952-66
9	32	61.5	10	4	US-09-521-650-66
10	32	61.5	10	4	US-09-168-888-66
11	32	61.5	102	2	US-08-580-988A-23
12	32	61.5	132	2	US-08-460-694-4
13	32	61.5	152	3	US-08-460-744-4
14	32	61.5	152	3	US-07-667-711B-4
15	32	61.5	173	1	US-08-193-977-7
16	32	61.5	189	2	US-08-464-517-21
17	32	61.5	189	2	US-08-246-361A-21
18	32	61.5	189	3	US-08-463-772-21
19	32	61.5	189	5	PCT-US93-05000-21
20	32	61.5	236	2	US-08-464-517-22
21	32	61.5	236	2	US-08-246-361A-22
22	32	61.5	236	5	PCT-US93-05000-22
23	32	61.5	280	3	US-08-464-517-6
24	32	61.5	280	3	US-08-463-772-6
25	32	61.5	289	2	US-08-246-361A-4
26	32	61.5	289	5	PCT-US93-05000-4
27	32	61.5	289	5	PCT-US93-05000-4

28	32	61.5	291	5	PCT-US93-05000-6	Sequence 6, Appl
29	32	61.5	292	2	US-08-464-517-23	Sequence 23, Appl
30	32	61.5	292	2	US-08-246-361A-6	Sequence 6, Appl
31	32	61.5	292	2	US-08-246-361A-23	Sequence 23, Appl
32	32	61.5	292	2	US-08-463-772-23	Sequence 23, Appl
33	32	61.5	292	5	PCT-US93-05000-23	Sequence 23, Appl
34	32	61.5	295	1	US-07-947-120-8	Sequence 8, Appl
35	32	61.5	295	1	US-08-472-893A-8	Sequence 8, Appl
36	32	61.5	295	1	US-08-460-694-2	Sequence 2, Appl
37	32	61.5	295	2	US-08-464-517-19	Sequence 19, Appl
38	32	61.5	295	2	US-08-464-517-20	Sequence 20, Appl
39	32	61.5	295	2	US-08-246-361A-19	Sequence 19, Appl
40	32	61.5	295	2	US-08-246-361A-20	Sequence 20, Appl
41	32	61.5	295	3	US-08-463-772-19	Sequence 19, Appl
42	32	61.5	295	3	US-08-463-772-20	Sequence 20, Appl
43	32	61.5	295	3	US-08-460-744-2	Sequence 2, Appl
44	32	61.5	295	3	US-07-667-711B-2	Sequence 2, Appl
45	32	61.5	295	3	US-08-947-492-8	Sequence 8, Appl
46	32	61.5	295	5	PCT-US93-05000-2	Sequence 2, Appl
47	32	61.5	295	5	PCT-US93-05000-19	Sequence 19, Appl
48	32	61.5	295	5	PCT-US93-05000-20	Sequence 20, Appl
49	32	61.5	309	3	US-08-464-517-4	Sequence 4, Appl
50	32	61.5	309	3	US-08-463-772-4	Sequence 4, Appl
51	32	61.5	529	4	US-09-240-639-4	Sequence 4, Appl
52	32	61.5	618	2	US-08-770-761A-3	Sequence 3, Appl
53	32	61.5	647	2	US-08-770-761A-8	Sequence 8, Appl
54	32	61.5	660	2	US-08-770-761A-2	Sequence 2, Appl
55	32	61.5	662	2	US-08-770-761A-5	Sequence 5, Appl
56	32	61.5	705	2	US-08-770-761A-7	Sequence 7, Appl
57	32	61.5	819	2	US-08-464-517-7	Sequence 7, Appl
58	32	61.5	819	2	US-08-246-361A-7	Sequence 7, Appl
59	32	61.5	819	3	US-08-463-772-7	Sequence 7, Appl
60	32	61.5	819	5	PCT-US93-05000-7	Sequence 7, Appl
61	31	59.6	59	4	US-08-963-851-14	Sequence 14, Appl
62	31	59.6	622	2	US-08-459-146-2	Sequence 2, Appl
63	31	59.6	622	2	US-08-459-065-2	Sequence 2, Appl
64	30	57.7	13	4	US-09-288-391-23	Sequence 23, Appl
65	30	57.7	13	4	US-09-288-391-22	Sequence 22, Appl
66	30	57.7	117	4	US-08-444-818-44	Sequence 44, Appl
67	30	57.7	121	4	US-09-152-060-68	Sequence 68, Appl
68	30	57.7	121	4	US-09-152-060-85	Sequence 85, Appl
69	30	57.7	122	2	US-08-879-995A-1	Sequence 1, Appl
70	30	57.7	122	3	US-09-215-096-1	Sequence 1, Appl
71	30	57.7	132	4	US-08-444-818-52	Sequence 52, Appl
72	30	57.7	159	2	US-08-844-086-4	Sequence 4, Appl
73	30	57.7	159	3	US-09-018-211-4	Sequence 4, Appl
74	30	57.7	211	5	PCT-US94-04174-18	Sequence 18, Appl
75	30	57.7	241	3	US-08-834-776A-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1  
US-09-228-986-73  
; Sequence 73, Application US/09228986  
; Patent No. 6359198  
; GENERAL INFORMATION:  
; APPLICANT: Strabala, Timothy  
; APPLICANT: Nieuwenhuizen, Niels  
; TITLE OF INVENTION: Compositions Isolated from Plant Cells  
and Their Use in the Modification of Plant Cell Signalling  
; FILE REFERENCE: 11000/1020  
; CURRENT APPLICATION NUMBER: US/09/228.986  
; CURRENT FILING DATE: 1999-01-12  
; NUMBER OF SEQ ID NOS: 130  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 73  
; LENGTH: 947  
; TYPE: PRT  
; ORGANISM: Pinus radiata  
US-09-228-986-73

2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
841

NO  
-08-637-759B-236

# CONCLUSIONS

ERROR APPLICATION DATA:  
APPLICATION NUMBER: 08/627 750



us-09-909-164-10.rai

Wed Jun 11 15:41:53 2003

QY 1 EEVVPXGMSYS 11  
:|: | | | |  
Db 399 KEICPGMGYT 409

RESULT 7  
5177197-30  
; Patent No. 5177197  
; APPLICANT: KANZAKI, TETSUTO; OLOFSSON, ANDERS; MOREN, ANITA;  
; WERNSTEDT, CHRISTER; HELLMAN, ULF; MIYAZONO, KOHEI; CLAESSEN-WELSH,  
; LENA; HELDIN, CARL-HENRIK  
; TITLE OF INVENTION: ISOLATED NUCLEOTIDE SEQUENCE EXPRESSING  
; HUMAN TRANSFORMING GROWTH FACTOR-BETAL-BINDING PROTEIN  
; NUMBER OF SEQUENCES: 53  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/487,343  
; FILING DATE: 27-FEB-1990  
; SEQ ID NO: 30:  
; LENGTH: 1394  
5177197-30

Query Match 63.5%; Score 33; DB 6; Length 1394;  
Best Local Similarity 45.5%; Pred. No. 2.6e+02;  
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSYS 11  
:|: | | | |  
Db 399 KEICPGMGYT 409

RESULT 8  
US-09-357-952-66  
; Sequence 66, Application US/09357952  
; Patent No. 6248904  
; GENERAL INFORMATION:  
; APPLICANT: Zhang, Han-Zhong  
; APPLICANT: Cai, Sui Xiong  
; APPLICANT: Drewe, John A.  
; APPLICANT: Yang, Wu  
; TITLE OF INVENTION: No. 6248904el Fluorescence Dyes and Their Applications for W  
; TITLE OF INVENTION: Fluorescence Screening Assays for Caspases, Peptidases, Prot  
; TITLE OF INVENTION: Other Enzymes and the Use Thereof  
; FILE REFERENCE: 1735.0030001  
; CURRENT APPLICATION NUMBER: US/09/357,952  
; CURRENT FILING DATE: 1999-07-21  
; EARLIER APPLICATION NUMBER: US 60/093,642  
; EARLIER FILING DATE: 21-JUL-1998  
; NUMBER OF SEQ ID NOS: 139  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 66  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
; OTHER INFORMATION: Peptide  
US-09-357-952-66

Query Match 61.5%; Score 32; DB 4; Length 10;  
Best Local Similarity 50.0%; Pred. No. 1.5;  
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10  
:|: | | | |  
Db 1 DDIVPCMSY 10

RESULT 9  
US-09-521-650-66  
; Sequence 66, Application US/09521650  
; Patent No. 6335429  
; GENERAL INFORMATION:  
; APPLICANT: Weber, Eckard

Query Match 63.5%; Score 33; DB 4; Length 45;  
Best Local Similarity 60.0%; Pred. No. 5.1;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10  
:|: | | | |  
Db 1 EEISPLGWSY 10

RESULT 5  
5177197-51  
; Patent No. 5177197  
; APPLICANT: KANZAKI, TETSUTO; OLOFSSON, ANDERS; MOREN, ANITA;  
; WERNSTEDT, CHRISTER; HELLMAN, ULF; MIYAZONO, KOHEI; CLAESSEN-WELSH,  
; LENA; HELDIN, CARL-HENRIK  
; TITLE OF INVENTION: ISOLATED NUCLEOTIDE SEQUENCE EXPRESSING  
; HUMAN TRANSFORMING GROWTH FACTOR-BETAL-BINDING PROTEIN  
; NUMBER OF SEQUENCES: 53  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/487,343  
; FILING DATE: 27-FEB-1990  
; SEQ ID NO: 51:  
; LENGTH: 65  
5177197-51

Query Match 63.5%; Score 33; DB 6; Length 65;  
Best Local Similarity 45.5%; Pred. No. 7.8;  
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSYS 11  
:|: | | | |  
Db 52 KEICPGMGYT 62

RESULT 6  
5177197-1  
; Patent No. 5177197  
; APPLICANT: KANZAKI, TETSUTO; OLOFSSON, ANDERS; MOREN, ANITA;  
; WERNSTEDT, CHRISTER; HELLMAN, ULF; MIYAZONO, KOHEI; CLAESSEN-WELSH,  
; LENA; HELDIN, CARL-HENRIK  
; TITLE OF INVENTION: ISOLATED NUCLEOTIDE SEQUENCE EXPRESSING  
; HUMAN TRANSFORMING GROWTH FACTOR-BETAL-BINDING PROTEIN  
; NUMBER OF SEQUENCES: 53  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/487,343  
; FILING DATE: 27-FEB-1990  
; SEQ ID NO: 1:  
; LENGTH: 410  
5177197-1

Query Match 63.5%; Score 33; DB 6; Length 410;  
Best Local Similarity 45.5%; Pred. No. 63;  
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

;; APPLICANT: Cai, Sui Xiong  
;; APPLICANT: Keana, John F.W.  
;; APPLICANT: Drewe, John A.  
;; APPLICANT: Zhang, Han-Zhong  
;; TITLE OF INVENTION: No. 6335429el Fluorogenic or Fluorescent Reporter Molecules and  
;; TITLE OF INVENTION: Their Applications for Whole-Cell Fluorescence  
;; TITLE OF INVENTION: Screening Assays for Caspases and Other Enzymes and the  
;; TITLE OF INVENTION: Use Thereof  
;; FILE REFERENCE: 1735.0290002  
;; CURRENT APPLICATION NUMBER: US/09/521.650  
;; CURRENT FILING DATE: 2000-03-08  
;; EARLIER APPLICATION NUMBER: 09/168,888  
;; EARLIER FILING DATE: 1998-10-09  
;; EARLIER APPLICATION NUMBER: US 60/061,582  
;; EARLIER FILING DATE: 1997-10-10  
;; EARLIER APPLICATION NUMBER: US 09/033,661  
;; EARLIER FILING DATE: 1998-03-03  
;; NUMBER OF SEQ ID NOS: 142  
;; SOFTWARE: PatentIn Ver. 2.0  
;; SEQ ID NO 66  
;; LENGTH: 10  
;; TYPE: PRT  
;; ORGANISM: Artificial Sequence  
;; FEATURE:  
;; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-09-521-650-66

Query Match 61.5%; Score 32; DB 4; Length 10;  
Best Local Similarity 50.0%; Pred. No. 1.5;  
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10  
Db 1 DDIVPCMSY 10

RESULT 10  
US-09-168-888-66  
; Sequence 66, Application US/09168888  
; Patent No. 6342611  
; GENERAL INFORMATION:  
; APPLICANT: Weber, Eckard  
; APPLICANT: Cai, Sui Xiong  
; APPLICANT: Keana, John F.W.  
; APPLICANT: Drewe, John A.  
; APPLICANT: Zhang, Han-Zhong  
; TITLE OF INVENTION: No. 6342611el Fluorogenic or Fluorescent Reporter Molecules and  
; TITLE OF INVENTION: Their Applications for Whole-Cell Fluorescence  
; TITLE OF INVENTION: Screening Assays for Caspases and Other Enzymes and the  
; FILE REFERENCE: 1735.0290002  
; CURRENT APPLICATION NUMBER: US/09/168,888  
; CURRENT FILING DATE: 1998-10-09  
; EARLIER APPLICATION NUMBER: US 60/061,582  
; EARLIER FILING DATE: 1997-10-10  
; EARLIER APPLICATION NUMBER: US 09/033,661  
; EARLIER FILING DATE: 1998-03-03  
; NUMBER OF SEQ ID NOS: 142  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 66  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-09-168-888-66

Query Match 61.5%; Score 32; DB 4; Length 10;  
Best Local Similarity 50.0%; Pred. No. 1.5;  
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10  
Db 1 DDIVPCMSY 10

RESULT 11  
US-08-580-988A-23  
; Sequence 23, Application US/08580988A  
; Patent No. 5856161  
; GENERAL INFORMATION:  
; APPLICANT: Aggarwal et al.  
; TITLE OF INVENTION: Tumor Necrosis Factor  
; TITLE OF INVENTION: Receptor-I-Associated Protein Kinase And Methods  
; TITLE OF INVENTION: For Its Use  
; NUMBER OF SEQUENCES: 27  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dr. Benjamin A. Adler  
; STREET: 8011 Candle Lane  
; CITY: Houston  
; STATE: Texas  
; COUNTRY: USA  
; ZIP: 77071  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 1.44 Mb floppy disk  
; COMPUTER: Apple Macintosh  
; OPERATING SYSTEM: Macintosh  
; SOFTWARE: Microsoft Word for Macintosh  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/580,988A  
; FILING DATE: January 3, 1996  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Benjamin Aaron Adler, Ph.D., J.D.  
; REGISTRATION NUMBER: 35,423  
; REFERENCE/DOCKET NUMBER: D5721CIP2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 713-777-2321  
; TELEFAX: 713-777-6908  
; INFORMATION FOR SEQ ID NO: 23:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 102 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE:  
; DESCRIPTION: protein  
; HYPOTHETICAL: no  
; ANTI-SENSE: no  
; FRAGMENT TYPE: internal  
; ORIGINAL SOURCE:  
US-08-580-988A-23

Query Match 61.5%; Score 32; DB 2; Length 102;  
Best Local Similarity 60.0%; Pred. No. 21;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10  
Db 24 EEVFLAMNY 33

RESULT 12  
US-08-460-694-4  
; Sequence 4, Application US/08460694  
; Patent No. 5858655  
; GENERAL INFORMATION:  
; APPLICANT: Arnold, Andrew  
; TITLE OF INVENTION: PRAD1 Cyclin and its cDNA  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:

us-09-909-164-10-rai

Wed Jun 11 15:41:53 2003

ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
STREET: 1100 New York Avenue, N.W., Suite 600  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/460,694  
FILING DATE: 02-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: McConathy, Evelyn H.  
REGISTRATION NUMBER: 35,279  
REFERENCE/DOCKET NUMBER: 0609.4070002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-2600  
TELEFAX: 202-371-2540  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 152 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-460-694-4

Query Match 61.5%; Score 32; DB 2; Length 152;  
Best Local Similarity 60.0%; Pred. No. 33;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
QY 1 EEVVPXGMSY 10  
Db 20 EEVFPPLAMNY 29

RESULT 13  
US-08-460-744-4  
Sequence 4, Application US/08460744  
Patent No. 6107541  
GENERAL INFORMATION:  
APPLICANT: Arnold, Andrew  
TITLE OF INVENTION: Prad1 Cyclin and its cdna  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
STREET: 1100 New York Avenue, N.W., Suite 600  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/460,744  
FILING DATE: 02-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: McConathy, Evelyn H.  
REGISTRATION NUMBER: 35,279  
REFERENCE/DOCKET NUMBER: 0609.4070005  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-2600  
TELEFAX: 202-371-2540  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:

LENGTH: 152 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-460-744-4  
Query Match 61.5%; Score 32; DB 3; Length 152;  
Best Local Similarity 60.0%; Pred. No. 33;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
QY 1 EEVVPXGMSY 10  
Db 20 EEVFPPLAMNY 29

RESULT 14  
US-07-667-711B-4  
Sequence 4, Application US/07667711B  
Patent No. 6110700  
GENERAL INFORMATION:  
APPLICANT: ARNOLD, ANDREW  
TITLE OF INVENTION: Prad1 Cyclin and its cdna  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
STREET: 1100 NEW YORK AVE., NW, SUITE 600  
CITY: WASHINGTON  
STATE: DC  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/667,711B  
FILING DATE: 11-MAR-1991  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: MCPHAIL, DONALD R.  
REGISTRATION NUMBER: 35,811  
REFERENCE/DOCKET NUMBER: 0609.4070000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 371-2600  
TELEFAX: (202) 371-2540  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 152 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: not relevant  
MOLECULE TYPE: peptide  
US-07-667-711B-4

Query Match 61.5%; Score 32; DB 3; Length 152;  
Best Local Similarity 60.0%; Pred. No. 33;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
QY 1 EEVVPXGMSY 10  
Db 20 EEVFPPLAMNY 29

RESULT 15  
US-08-193-977-7  
Sequence 7, Application US/08193977  
Patent No. 5625031  
GENERAL INFORMATION:  
APPLICANT: WEBSTER, KEVIN R.  
APPLICANT: COLEMAN, KEVIN G.  
TITLE OF INVENTION: PEPTIDE INHIBITORS OF THE P33CDK2 AND

;; TITLE OF INVENTION: P34CDC2 CELL CYCLE REGULATORY KINASES AND HUMAN  
;; NUMBER OF SEQUENCES: 34  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: REED & ROBINS  
;; STREET: 635 BRYANT STREET  
;; CITY: PALO ALTO  
;; STATE: CALIFORNIA  
;; COUNTRY: UNITED STATES OF AMERICA  
;; ZIP: 94301  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: Patentin Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/193,977  
;; FILING DATE: 08-FEB-1994  
;; CLASSIFICATION: 530  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: ROBINS, ROBERTA L.  
;; REGISTRATION NUMBER: 33,208  
;; REFERENCE/DOCKET NUMBER: 5998-0016  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (415) 617-8999  
;; TELEFAX: (415) 327-3231  
;; INFORMATION FOR SEQ ID NO: 7:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 173 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
US-08-193-977-7

Query Match 61.5%; Score 32; DB 1; Length 173;  
Best Local Similarity 60.0%; Pred. No. 38;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10  
||| | | |  
Db 55 EEVFFPLAMNY 64

RESULT 16  
US-08-464-517-21  
; Sequence 21, Application US/08464517  
; Patent No. 5869640  
; GENERAL INFORMATION:  
; APPLICANT: BEACH, David H.  
; TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO  
; NUMBER OF SEQUENCES: 50  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD  
; STREET: 60 State Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII(text)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/464,517  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/963,308  
; FILING DATE: 16-OCT-1992  
; APPLICATION NUMBER: US 07/888,178  
; FILING DATE: 26-MAY-1992

;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/701,514  
;; FILING DATE: 16-MAY-1991  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Matthew P. Vincent  
;; REGISTRATION NUMBER: 36,709  
;; REFERENCE/DOCKET NUMBER: MII-004C  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (617) 227-7400  
;; TELEFAX: (617) 227-5941  
;; INFORMATION FOR SEQ ID NO: 21:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 189 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
US-08-464-517-21

Query Match 61.5%; Score 32; DB 2; Length 189;  
Best Local Similarity 60.0%; Pred. No. 42;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10  
||| | | |  
Db 74 EEVFFPLAMNY 83

RESULT 17  
US-08-246-361A-21  
; Sequence 21, Application US/08246361A  
; Patent No. 5998582  
; GENERAL INFORMATION:  
; APPLICANT: BEACH, David H.  
; TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO  
; NUMBER OF SEQUENCES: 50  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD  
; STREET: 60 State Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII(text)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/246,361A  
; FILING DATE: 19-MAY-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/963,308  
; FILING DATE: 16-OCT-1992  
; APPLICATION NUMBER: US 07/888,178  
; FILING DATE: 26-MAY-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/701,514  
; FILING DATE: 16-MAY-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Matthew P. Vincent  
; REGISTRATION NUMBER: 36,709  
; REFERENCE/DOCKET NUMBER: MII-004C  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 227-7400  
; TELEFAX: (617) 227-5941  
; INFORMATION FOR SEQ ID NO: 21:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 189 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single

us-09-909-164-10-rai

Wed Jun 11 15:41:53 2003

```

; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-246-361A-21

Query Match 61.5%; Score 32; DB 2; Length 189;
Best Local Similarity 60.0%; Pred. No. 42;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10
DB 74 EEVFLAMNY 83

RESULT 18
US-08-463-772-21
; Sequence 21, Application US/08463772
; Patent No. 6066501
; GENERAL INFORMATION:
; APPLICANT: BEACH, David H.
; TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII(text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/463,772
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/963,308
; FILING DATE: 16-OCT-1992
; APPLICATION NUMBER: US 07/888,178
; FILING DATE: 26-MAY-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/701,514
; FILING DATE: 16-MAY-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Matthew P. Vincent
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: MII-004C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 189 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-463-772-21

Query Match 61.5%; Score 32; DB 3; Length 189;
Best Local Similarity 60.0%; Pred. No. 42;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10
DB 74 EEVFLAMNY 83

RESULT 19
PCT-US93-05000-21
; Sequence 21, Application PC/TUS9305000
; APPLICANT: MITOXIX
; TITLE OF INVENTION: D-Type Cyclin and Uses Related Thereto
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: US
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/05000
; FILING DATE: 19930525
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/888,178
; FILING DATE: 26-MAY-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: CSHL91-02A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-6240
; TELEFAX: 616-861-9540
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 189 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; PCT-US93-05000-21

Query Match 61.5%; Score 32; DB 5; Length 189;
Best Local Similarity 60.0%; Pred. No. 42;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10
DB 74 EEVFLAMNY 83

RESULT 20
US-08-464-517-22
; Sequence 22, Application US/08464517
; Patent No. 5869640
; GENERAL INFORMATION:
; APPLICANT: BEACH, David H.
; TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII(text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/464,517
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/963,308
```

```

; FILING DATE: 16-OCT-1992
; APPLICATION NUMBER: US 07/888,178
; FILING DATE: 26-MAY-1992
; PRIOR APPLICATION DATA: US 07/701,514
; FILING DATE: 16-MAY-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Matthew P. Vincent
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: MII-004C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-7400
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 236 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-464-517-22

```

```

Query Match 61.5%; Score 32; DB 2; Length 236;
Best Local Similarity 60.0%; Pred. No. 54;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

```

```

QY 1 EEVVPXGMSY 10
   ||| | | |
Db 20 EEVFPPLAMNY 29

```

## RESULT 21

```

US-08-246-361A-22
; Sequence 22, Application US/08246361A
; Patent No. 5998582
; GENERAL INFORMATION:
; APPLICANT: BEACH, David H.
; TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII(text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/246,361A
; FILING DATE: 19-MAY-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/963,308
; FILING DATE: 16-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/888,178
; FILING DATE: 26-MAY-1992
; APPLICATION DATA:
; FILING DATE: 16-MAY-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Matthew P. Vincent
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: MII-004C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-7400
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:

```

```

; LENGTH: 236 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-246-361A-22

```

```

Query Match 61.5%; Score 32; DB 2; Length 236;
Best Local Similarity 60.0%; Pred. No. 54;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

```

```

QY 1 EEVVPXGMSY 10
   ||| | | |
Db 20 EEVFPPLAMNY 29

```

## RESULT 22

```

US-08-463-772-22
; Sequence 22, Application US/08463772
; Patent No. 6066501
; GENERAL INFORMATION:
; APPLICANT: BEACH, David H.
; TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII(text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/463,772
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/963,308
; FILING DATE: 16-OCT-1992
; APPLICATION NUMBER: US 07/888,178
; FILING DATE: 26-MAY-1992
; APPLICATION DATA:
; APPLICATION NUMBER: US 07/701,514
; FILING DATE: 16-MAY-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Matthew P. Vincent
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: MII-004C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 236 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-463-772-22

```

```

Query Match 61.5%; Score 32; DB 3; Length 236;
Best Local Similarity 60.0%; Pred. No. 54;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

```

```

QY 1 EEVVPXGMSY 10
   ||| | | |
Db 20 EEVFPPLAMNY 29

```

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

```

; SEQUENCE CHARACTERISTICS:  
; LENGTH: 280 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-463-772-6

Query Match 61.5%; Score 32; DB 3; Length 280;  
Best Local Similarity 60.0%; Pred. No. 65;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10  
||| | |:  
Db 75 EEVPLAMNY 84

Search completed: June 10, 2003, 13:51:33  
Job time : 9.64286 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2003, 13:46:50 ; Search time 15 Seconds  
(without alignments)  
75.710 Million cell updates/sec

Title: US-09-909-164-10  
Perfect score: 52  
Sequence: 1 EENVXGMSYS 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 392085 seqs, 103240269 residues

Total number of hits satisfying chosen parameters: 392085

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 75 summaries

Database :

Published Applications\_AA.\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pap.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pap.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pap.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pap.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pap.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pap.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pap.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pap.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pap.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pap.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pap.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pap.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pap.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pap.\*

pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	38	73.1	3472	9	US-10-027-806-4
2	38	73.1	3472	9	US-10-034-623-4
3	38	73.1	3472	9	US-10-027-801-4
4	34	65.4	947	9	US-10-101-464A-73
5	33	63.5	426	9	US-10-214-766-43
6	33	63.5	478	9	US-09-924-340-108
7	33	63.5	478	9	US-09-992-600A-108
8	33	63.5	478	9	US-09-746-783-184
9	33	63.5	478	9	US-10-000-489-108
10	33	63.5	478	9	US-10-000-986-108
11	33	63.5	653	9	US-09-820-843A-26
12	32	61.5	10	10	US-09-947-387-66
13	32	61.5	254	10	US-09-778-927A-53
14	32	61.5	289	9	US-10-024-056-2
15	32	61.5	289	9	US-10-024-066-4
16	32	61.5	289	10	US-09-919-497-54
17	32	61.5	295	10	US-09-925-300-1061
18	32	61.5	529	10	US-09-923-304-4
19	32	61.5	691	9	US-10-101-921-4

#### ALIGNMENTS

#### RESULT 1

US-10-027-806-4

; Sequence 4, Application US/10027806  
; Patent No. US20020160476A1

; GENERAL INFORMATION:

; APPLICANT: Swanson, Ronald V.

; APPLICANT: Feldman, Robert A.

; APPLICANT: Schleper, Christa

; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM CENARCHAEUM SYMBIOSUM

; FILE REFERENCE: DCCP.002A

; CURRENT APPLICATION NUMBER: US/10/027,806

; CURRENT FILING DATE: 2001-12-21

20	32	61.5	691	10	US-09-925-731-2	Sequence 2, Appli
21	32	61.5	1377	10	US-09-815-242-10384	Sequence 10384, A
22	32	61.5	2799	9	US-10-151-736-4	Sequence 4, Appli
23	31	59.6	53	9	US-10-092-154-878	Sequence 878, App
24	31	59.6	53	10	US-09-764-847-878	Sequence 878, App
25	31	59.6	59	10	US-09-948-080-14	Sequence 14, Appl
26	31	59.6	161	9	US-09-738-626-5124	Sequence 5124, Ap
27	31	59.6	163	9	US-10-117-846-20	Sequence 20, Appl
28	31	59.6	192	9	US-09-986-480-171	Sequence 171, App
29	31	59.6	198	10	US-09-731-872-334	Sequence 334, App
30	31	59.6	223	9	US-09-738-626-6349	Sequence 6349, Ap
31	31	59.6	299	10	US-09-815-242-10697	Sequence 10697, A
32	31	59.6	381	9	US-09-975-139-5	Sequence 5, Appli
33	31	59.6	702	9	US-10-280-403-2	Sequence 2, Appli
34	31	59.6	702	10	US-09-907-479-2	Sequence 4454, Ap
35	31	59.6	763	9	US-09-738-626-4454	Sequence 5136, Ap
36	31	59.6	1053	10	US-09-815-242-5136	Sequence 5136, Ap
37	31	59.6	1407	10	US-09-815-242-10439	Sequence 10439, A
38	31	59.6	1426	10	US-09-912-020-340	Sequence 340, App
39	30	57.7	7	9	US-09-909-062-1	Sequence 1, Appli
40	30	57.7	7	9	US-09-909-062-9	Sequence 9, Appli
41	30	57.7	7	9	US-09-909-062-130	Sequence 68, Appl
42	30	57.7	121	9	US-09-852-797-68	Sequence 85, Appl
43	30	57.7	121	9	US-09-852-797-85	Sequence 85, Appl
44	30	57.7	121	10	US-09-853-161-68	Sequence 68, Appl
45	30	57.7	121	10	US-09-853-161-85	Sequence 85, Appl
46	30	57.7	121	10	US-09-852-659A-68	Sequence 68, Appl
47	30	57.7	121	10	US-09-852-659A-85	Sequence 85, Appl
48	30	57.7	135	9	US-09-992-598-359	Sequence 359, App
49	30	57.7	135	9	US-09-989-293A-359	Sequence 359, App
50	30	57.7	135	9	US-09-989-735-359	Sequence 359, App
51	30	57.7	135	9	US-09-990-444-359	Sequence 359, App
52	30	57.7	135	9	US-09-989-730-359	Sequence 359, App
53	30	57.7	135	9	US-09-990-436-359	Sequence 359, App
54	30	57.7	135	9	US-09-991-181-359	Sequence 359, App
55	30	57.7	135	9	US-09-983-687-359	Sequence 359, App
56	30	57.7	135	9	US-09-989-734-359	Sequence 359, App
57	30	57.7	135	9	US-09-997-653-359	Sequence 359, App
58	30	57.7	135	9	US-10-174-590-444	Sequence 444, App
59	30	57.7	135	9	US-10-176-758-444	Sequence 444, App
60	30	57.7	135	9	US-10-175-737-444	Sequence 444, App
61	30	57.7	135	9	US-09-993-667-359	Sequence 359, App
62	30	57.7	135	9	US-10-173-706-444	Sequence 444, App
63	30	57.7	135	9	US-10-175-728-444	Sequence 444, App
64	30	57.7	135	9	US-10-175-752-444	Sequence 444, App
65	30	57.7	135	9	US-10-176-482-444	Sequence 444, App
66	30	57.7	135	9	US-10-176-757-444	Sequence 444, App
67	30	57.7	135	9	US-10-176-913-444	Sequence 444, App
68	30	57.7	135	9	US-10-180-552-444	Sequence 444, App
69	30	57.7	135	9	US-10-180-557-444	Sequence 444, App
70	30	57.7	135	9	US-09-990-438-359	Sequence 359, App
71	30	57.7	135	9	US-09-990-562-359	Sequence 359, App
72	30	57.7	135	9	US-09-997-428-359	Sequence 359, App
73	30	57.7	135	9	US-09-997-666-359	Sequence 359, App
74	30	57.7	135	9	US-10-173-700-444	Sequence 444, App
75	30	57.7	135	9	US-10-174-572-444	Sequence 444, App

; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/408,020  
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-09-29  
; NUMBER OF SEQ ID NOS: 123  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 4  
; LENGTH: 3472  
; TYPE: PRT  
; ORGANISM: Cenarchaeum symbiosum  
US-10-027-806-4

Query Match 73.1%; Score 38; DB 9; Length 3472;  
Best Local Similarity 54.5%; Pred. No. 1.2e+02;  
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVPXGMSYS 11  
Db 2294 EDVPRGISFS 2304  
|:|:| |:|:|

RESULT 2  
US-10-034-623-4  
; Sequence 4, Application US/10034623  
; Publication No. US20020198365A1  
; GENERAL INFORMATION:  
; APPLICANT: Swanson, Ronald V.  
; APPLICANT: Feldman, Robert A.  
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM CENARCHAEUM SYMBIOSUM  
; FILE REFERENCE: DCOEP.002A  
; CURRENT APPLICATION NUMBER: US/10/034,623  
; PRIOR FILING DATE: 2001-12-21  
; PRIOR APPLICATION NUMBER: 09/408,020  
; PRIOR FILING DATE: 1999-09-29  
; PRIOR APPLICATION NUMBER: 60/102,294  
; PRIOR FILING DATE: 1998-09-29  
; NUMBER OF SEQ ID NOS: 123  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 4  
; LENGTH: 3472  
; TYPE: PRT  
; ORGANISM: Cenarchaeum symbiosum  
US-10-034-623-4

Query Match 73.1%; Score 38; DB 9; Length 3472;  
Best Local Similarity 54.5%; Pred. No. 1.2e+02;  
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVPXGMSYS 11  
Db 2294 EDVPRGISFS 2304  
|:|:| |:|:|

RESULT 3  
US-10-027-801-4  
; Sequence 4, Application US/10027801  
; Publication No. US20030054364A1  
; GENERAL INFORMATION:  
; APPLICANT: Swanson, Ronald V.  
; APPLICANT: Feldman, Robert A.  
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM CENARCHAEUM SYMBIOSUM  
; FILE REFERENCE: DCOEP.002A  
; CURRENT APPLICATION NUMBER: US/10/027,801  
; CURRENT FILING DATE: 2001-12-21  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/408,020  
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-09-29  
; NUMBER OF SEQ ID NOS: 123  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 4  
; LENGTH: 3472  
; TYPE: PRT  
; ORGANISM: Cenarchaeum symbiosum  
US-10-027-801-4

Query Match 73.1%; Score 38; DB 9; Length 3472;  
Best Local Similarity 54.5%; Pred. No. 1.2e+02;  
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVPXGMSYS 11  
Db 2294 EDVPRGISFS 2304  
|:|:| |:|:|

RESULT 4  
US-10-101-464A-73  
; Sequence 73, Application US/10101464A  
; Publication No. US20030046728A1  
; GENERAL INFORMATION:  
; APPLICANT: Strabala, Timothy  
; APPLICANT: Nieuwenhuizen, Nicolaas  
; TITLE OF INVENTION: Compositions Isolated from Plant Cells  
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signaling  
; FILE REFERENCE: 11000.1020C2  
; CURRENT APPLICATION NUMBER: US/10/101.464A  
; CURRENT FILING DATE: 2002-03-18  
; PRIOR APPLICATION NUMBER: 09/704,302  
; PRIOR FILING DATE: 2000-11-01  
; PRIOR APPLICATION NUMBER: 09/228,986  
; PRIOR FILING DATE: 1999-01-12  
; PRIOR APPLICATION NUMBER: 60/162,866  
; PRIOR FILING DATE: 1999-11-01  
; PRIOR APPLICATION NUMBER: PCT/US00/00724  
; PRIOR FILING DATE: 2000-01-11  
; NUMBER OF SEQ ID NOS: 989  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 73  
; LENGTH: 947  
; TYPE: PRT  
; ORGANISM: Pinus radiata  
US-10-101-464A-73

Query Match 65.4%; Score 34; DB 9; Length 947;  
Best Local Similarity 66.7%; Pred. No. 1.8e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMSYS 11  
Db 686 VMPSGISYS 694  
|:|:| |:|:|

RESULT 5  
US-10-214-766-43  
; Sequence 43, Application US/10214766  
; Publication No. US20030084473A1  
; GENERAL INFORMATION:  
; APPLICANT: Gocal, Greg  
; TITLE OF INVENTION: NON-TRANSGENIC HERBICIDE RESISTANT PLANTS  
; FILE REFERENCE: CALL38  
; CURRENT APPLICATION NUMBER: US/10/214,766  
; CURRENT FILING DATE: 2002-08-09  
; PRIOR APPLICATION NUMBER: US 60/311,734  
; PRIOR FILING DATE: 2001-08-09  
; NUMBER OF SEQ ID NOS: 53  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 43  
; LENGTH: 426  
; TYPE: PRT  
; ORGANISM: Vibrio cholerae  
US-10-214-766-43

Query Match 63.5%; Score 33; DB 9; Length 426;  
Best Local Similarity 60.0%; Pred. No. 1.2e+02;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVPXGMSYS 10

Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 EVVPXGMSYS 11  
|||  
Db 239 EVAPAGASYN 248

## RESULT 8

US-09-746-783-184  
; Sequence 184, Application US/09746783  
; Publication No. US20030044935A1  
; GENERAL INFORMATION:  
; APPLICANT: Jacobs, Kenneth  
; McCOY, John M.  
; Lavallie, Edward R.  
; Racie, Lisa A.  
; Treacy, Maurice  
; Spaulding, Vikki  
; Agostino, Michael J.  
; Howes, Steven H.  
; Fechtel, Kim  
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES  
; ENCODING THEM  
; NUMBER OF SEQUENCES: 231  
; CORRESPONDENCE ADDRESS:  
; ADDRESSER: Genetics Institute, Inc.  
; STREET: 87 CambridgePark Drive  
; CITY: Cambridge  
; STATE: MA  
; COUNTRY: U.S.A.  
; ZIP: 02140  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/746.783  
; FILING DATE: 21-Dec-2000  
; CLASSIFICATION: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Milasincic, Debra J.  
; REGISTRATION NUMBER: 46,931  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 227-7400  
; TELEFAX: (617) 742-4214  
; INFORMATION FOR SEQ ID NO: 184:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 478 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: <Unknown>  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 184:  
US-09-746-783-184

Query Match 63.5%; Score 33; DB 9; Length 478;  
Best Local Similarity 60.0%; Pred. No. 1.4e+02;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 EVVPXGMSYS 11  
|||  
Db 239 EVAPAGASYN 248

## RESULT 9

US-10-000-489-108  
; Sequence 108, Application US/10000489  
; Publication No. US20030092011A1  
; GENERAL INFORMATION:  
; APPLICANT: Benjamin, Stephane  
; APPLICANT: Tanaka, Hiroaki  
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF

Db 223 EFVIPAGOSY 232  
|:|:|:|

## RESULT 6

US-09-924-340-108  
; Sequence 108, Application US/09924340  
; Publication No. US20030027248A1  
; GENERAL INFORMATION:  
; APPLICANT: Benjamin, Stephane  
; APPLICANT: Tanaka, Hiroaki  
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF  
; FILE REFERENCE: 91.US2.REG  
; CURRENT APPLICATION NUMBER: US/09/924,340  
; CURRENT FILING DATE: 2001-08-06  
; PRIOR APPLICATION NUMBER: US 60/305,456  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: US 60/302,277  
; PRIOR FILING DATE: 2001-06-29  
; PRIOR APPLICATION NUMBER: US 60/298,698  
; PRIOR FILING DATE: 2001-06-15  
; PRIOR APPLICATION NUMBER: US 60/293,574  
; PRIOR FILING DATE: 2001-05-25  
; NUMBER OF SEQ ID NOS: 112  
; SOFTWARE: JPatent  
; SEQ ID NO 108  
; LENGTH: 478  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-924-340-108

Query Match 63.5%; Score 33; DB 9; Length 478;  
Best Local Similarity 60.0%; Pred. No. 1.4e+02;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 EVVPXGMSYS 11  
|||  
Db 239 EVAPAGASYN 248

## RESULT 7

US-09-992-600A-108  
; Sequence 108, Application US/09992600A  
; Publication No. US20030027161A1  
; GENERAL INFORMATION:  
; APPLICANT: Benjamin, Stephane  
; APPLICANT: Tanaka, Hiroaki  
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF  
; FILE REFERENCE: 91.US4.DIV  
; CURRENT APPLICATION NUMBER: US/09/992,600A  
; CURRENT FILING DATE: 2001-11-13  
; PRIOR APPLICATION NUMBER: US 09/924,340  
; PRIOR FILING DATE: 2001-08-06  
; PRIOR APPLICATION NUMBER: PCT/IB01/01715  
; PRIOR FILING DATE: 2001-08-06  
; PRIOR APPLICATION NUMBER: US 60/305,456  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: US 60/302,277  
; PRIOR FILING DATE: 2001-06-29  
; PRIOR APPLICATION NUMBER: US 60/298,698  
; PRIOR FILING DATE: 2001-06-15  
; PRIOR APPLICATION NUMBER: US 60/293,574  
; PRIOR FILING DATE: 2001-05-25  
; NUMBER OF SEQ ID NOS: 114  
; SOFTWARE: JPatent  
; SEQ ID NO 108  
; LENGTH: 478  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-992-600A-108

Query Match 63.5%; Score 33; DB 9; Length 478;  
Best Local Similarity 60.0%; Pred. No. 1.4e+02;

FILE REFERENCE: 91.US6.DIV  
; CURRENT APPLICATION NUMBER: US/10/000,489  
; CURRENT FILING DATE: 2001-11-14  
; PRIOR APPLICATION NUMBER: US 09/924,340  
; PRIOR FILING DATE: 2001-08-06  
; PRIOR APPLICATION NUMBER: PCT/IB01/01715  
; PRIOR FILING DATE: 2001-08-06  
; PRIOR APPLICATION NUMBER: US 60/305,456  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: US 60/302,277  
; PRIOR FILING DATE: 2001-06-29  
; PRIOR APPLICATION NUMBER: US 60/298,698  
; PRIOR FILING DATE: 2001-06-15  
; PRIOR APPLICATION NUMBER: US 60/293,574  
; PRIOR FILING DATE: 2001-05-25  
; NUMBER OF SEQ ID NOS: 112  
; SOFTWARE: JPatent  
; SEQ ID NO 108  
; LENGTH: 478  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-000-489-108

Query Match 63.5%; Score 33; DB 9; Length 478;  
Best Local Similarity 60.0%; Pred. No. 1.4e+02;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 EVVVPXGMSYS 11  
Db 239 EVAPAGASYN 248  
||| | | |

## RESULT 10

US-10-000-986-108  
; Sequence 108, Application US/10000986  
; Publication No. US20030096247A1  
; GENERAL INFORMATION:  
; APPLICANT: Benjamin, Stephane  
; APPLICANT: Tanaka, Hiroaki  
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF  
; FILE REFERENCE: 91.US9.DIV  
; CURRENT APPLICATION NUMBER: US/10/000,986  
; CURRENT FILING DATE: 2001-11-14  
; PRIOR APPLICATION NUMBER: US 09/924,340  
; PRIOR FILING DATE: 2001-08-06  
; PRIOR APPLICATION NUMBER: PCT/IB01/01715  
; PRIOR FILING DATE: 2001-08-06  
; PRIOR APPLICATION NUMBER: US 60/305,456  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: US 60/302,277  
; PRIOR FILING DATE: 2001-06-29  
; PRIOR APPLICATION NUMBER: US 60/298,698  
; PRIOR FILING DATE: 2001-06-15  
; PRIOR APPLICATION NUMBER: US 60/293,574  
; PRIOR FILING DATE: 2001-05-25  
; NUMBER OF SEQ ID NOS: 112  
; SOFTWARE: JPatent  
; SEQ ID NO 108  
; LENGTH: 478  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-000-986-108

Query Match 63.5%; Score 33; DB 9; Length 478;  
Best Local Similarity 60.0%; Pred. No. 1.4e+02;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 EVVVPXGMSYS 11  
Db 239 EVAPAGASYN 248  
||| | | |

## RESULT 11

US-09-820-843A-26  
; Sequence 26, Application US/09820843A  
; Publication No. US20030039963A1  
; GENERAL INFORMATION:  
; APPLICANT: Council of Scientific and Industrial Research  
; TITLE OF INVENTION: A COMPUTATIONAL METHOD FOR THE IDENTIFICATION OF CANDIDATE P  
; TITLE OF INVENTION: USEFUL AS ANTI-INFECTIVES  
; FILE REFERENCE: Q63915  
; CURRENT APPLICATION NUMBER: US/09/820,843A  
; CURRENT FILING DATE: 2001-03-30  
; NUMBER OF SEQ ID NOS: 118  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 26  
; LENGTH: 653  
; TYPE: PRT  
; ORGANISM: Vibrio cholerae  
; NAME/KEY: misc\_feature  
; OTHER INFORMATION: iron(III) ABC transporter, permease protein  
; NAME/KEY: misc\_feature  
; OTHER INFORMATION: gi|9654609  
US-09-820-843A-26

Query Match 63.5%; Score 33; DB 9; Length 653;  
Best Local Similarity 66.7%; Pred. No. 1.9e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXGMS 9  
Db 300 EEVVPXGMSIT 308  
||||| | |

## RESULT 12

US-09-947-387-66  
; Sequence 66, Application US/09947387  
; Patent No. US20020150885A1  
; GENERAL INFORMATION:  
; APPLICANT: Weber, Eckard  
; APPLICANT: Cai, Sui Xiong  
; APPLICANT: Keana, John F.W.  
; APPLICANT: Drewe, John A.  
; APPLICANT: Zhang, Han-Zhong  
; TITLE OF INVENTION: No. US20020150885A1 Fluorogenic or Fluorescent Reporter Mol  
; TITLE OF INVENTION: Their Applications for Whole-Cell Fluorescence  
; TITLE OF INVENTION: Screening Assays for Caspases and Other Enzymes and the  
; TITLE OF INVENTION: Use Thereof  
; FILE REFERENCE: 1735.0290005  
; CURRENT APPLICATION NUMBER: US/09/947,387  
; CURRENT FILING DATE: 2001-09-07  
; PRIOR APPLICATION NUMBER: US 60/061,582  
; PRIOR FILING DATE: 1997-10-10  
; PRIOR APPLICATION NUMBER: US 60/145,746  
; PRIOR FILING DATE: 1998-03-03  
; PRIOR APPLICATION NUMBER: US 09/168,888  
; PRIOR FILING DATE: 1998-10-09  
; NUMBER OF SEQ ID NOS: 142  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 66  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
; OTHER INFORMATION: Peptide  
US-09-947-387-66

Query Match 61.5%; Score 32; DB 10; Length 10;  
Best Local Similarity 50.0%; Pred. No. 3.5;  
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10  
Db 1 DDIVPCMSY 10  
:::| | | |

us-09-909-164-10.rapb

Wed Jun 11 15:41:55 2003

```
; Sequence 4, Application US/10024066
; Patent No. US20020166134A1
; GENERAL INFORMATION:
; APPLICANT: Field, Loren J.
; TITLE OF INVENTION: PASUMARTHI, KISHORE BABU S.
; TITLE OF INVENTION: CARDIOMYOCYTES WITH ENHANCED PROLIFERATIVE POTENTIAL,
; TITLE OF INVENTION: AND METHODS FOR PREPARING AND USING SAME
; FILE REFERENCE: 7037-450
; CURRENT APPLICATION NUMBER: US/10/024,066
; CURRENT FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: 60/139,942
; PRIOR FILING DATE: 1999-06-18
; PRIOR APPLICATION NUMBER: PCT/US00/16827
; PRIOR FILING DATE: 2000-06-19
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 4
; LENGTH: 289
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-024-066-4
; Query Match 61.5%; Score 32; DB 9; Length 289;
; Best Local Similarity 60.0%; Pred. No. 1.3e+02;
; Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10
Db 74 EEVFPPLAMNY 83

RESULT 16
US-09-919-497-54
; Sequence 54, Application US/09919497
; Patent No. US20020106662A1
; GENERAL INFORMATION:
; APPLICANT: Mutter, George L.
; TITLE OF INVENTION: PROGNOSTIC CLASSIFICATION OF ENDOMETRIAL CANCER
; FILE REFERENCE: B0801/7225
; CURRENT APPLICATION NUMBER: US/09/919,497
; CURRENT FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: US 60/221,735
; PRIOR FILING DATE: 2000-07-31
; NUMBER OF SEQ ID NOS: 100
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 54
; LENGTH: 289
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-919-497-54
; Query Match 61.5%; Score 32; DB 10; Length 289;
; Best Local Similarity 60.0%; Pred. No. 1.3e+02;
; Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10
Db 74 EEVFPPLAMNY 83

RESULT 17
US-09-925-300-1061
; Sequence 1061, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PAL01
; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; US-09-925-300-1061
```

---

```
; Sequence 53, Application US/09778927A
; Patent No. US20020068342A1
; GENERAL INFORMATION:
; APPLICANT: KHOSRAVI, Rami et al.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID AND AMINO ACID SEQUENCES AND NOVEL
; TITLE OF INVENTION: VARIANTS OF ALTERNATIVE SPLICING
; FILE REFERENCE: 2786-0160P
; CURRENT APPLICATION NUMBER: US/09/778,927A
; CURRENT FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: IL 134453
; PRIOR FILING DATE: 2000-02-09
; PRIOR APPLICATION NUMBER: IL135341
; PRIOR FILING DATE: 2000-03-29
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 53
; LENGTH: 254
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)..(254)
; OTHER INFORMATION: Xaa = any amino acid, unknown, or other
; US-09-778-927A-53
; Query Match 61.5%; Score 32; DB 10; Length 254;
; Best Local Similarity 60.0%; Pred. No. 1.1e+02;
; Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10
Db 74 EEVFPPLAMNY 83

RESULT 14
US-10-024-066-2
; Sequence 2, Application US/10024066
; Patent No. US20020166134A1
; GENERAL INFORMATION:
; APPLICANT: Field, Loren J.
; TITLE OF INVENTION: PASUMARTHI, KISHORE BABU S.
; TITLE OF INVENTION: CARDIOMYOCYTES WITH ENHANCED PROLIFERATIVE POTENTIAL,
; TITLE OF INVENTION: AND METHODS FOR PREPARING AND USING SAME
; FILE REFERENCE: 7037-450
; CURRENT APPLICATION NUMBER: US/10/024,066
; CURRENT FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: 60/139,942
; PRIOR FILING DATE: 1999-06-18
; PRIOR APPLICATION NUMBER: PCT/US00/16827
; PRIOR FILING DATE: 2000-06-19
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 289
; TYPE: PRT
; ORGANISM: Mus musculus
; US-10-024-066-2
; Query Match 61.5%; Score 32; DB 9; Length 289;
; Best Local Similarity 60.0%; Pred. No. 1.3e+02;
; Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10
Db 73 EEVFPPLAMNY 82

RESULT 15
US-10-024-066-4
```

```

; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1061
; LENGTH: 295
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (243)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (277)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-300-1061

Query Match      61.5%; Score 32; DB 10; Length 295;
Best Local Similarity 60.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      2 EVVPXGMSYS 11
Db      52 EVLPTRMSYA 61
      ||| |||
      TYPE: PRT
      ORGANISM: Homo sapiens
      FEATURE:
      NAME/KEY: SITE
      LOCATION: (243)
      OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
      NAME/KEY: SITE
      LOCATION: (277)
      OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

US-09-923-304-4
; Sequence 4, Application US/09923304
; Patent No. US20020081612A1
; GENERAL INFORMATION:
; APPLICANT: KATZ, RUTH
; APPLICANT: JIANG, FENG
; TITLE OF INVENTION: DETECTION AND DIAGNOSIS OF SMOKING RELATED CANCERS
; FILE REFERENCE: UTSC:6580S
; CURRENT APPLICATION NUMBER: US/09/923,304
; CURRENT FILING DATE: 2001-08-06
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 529
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-923-304-4

Query Match      61.5%; Score 32; DB 10; Length 529;
Best Local Similarity 50.0%; Pred. No. 2.4e+02;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY      1 EEVVPXGMSY 10
Db      48 QEVLPPGLXY 57
      ||| |||
      TYPE: PRT
      ORGANISM: Homo sapiens
      FEATURE:
      NAME/KEY: SITE
      LOCATION: (243)
      OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
      NAME/KEY: SITE
      LOCATION: (277)
      OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

US-09-925-300-1061
; Sequence 4, Application US/10101921
; Publication No. US20030022199A1
; GENERAL INFORMATION:
; APPLICANT: Nezu, Jun-Ichi
; APPLICANT: Ose, Asuka
; APPLICANT: Tsuji, Akira
; TITLE OF INVENTION: TRANSPORTER GENES OATP-B, C, D, AND E
; FILE REFERENCE: 06501-104US1
; CURRENT APPLICATION NUMBER: US/10/101,921
; CURRENT FILING DATE: 2002-06-28
; PRIOR APPLICATION NUMBER: PCT/JP00/06416
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: JP 11/267835
; PRIOR FILING DATE: 1999-09-21
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4

```

```

; LENGTH: 691
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-101-921-4

Query Match      61.5%; Score 32; DB 9; Length 691;
Best Local Similarity 62.5%; Pred. No. 3.2e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      3 VVPXGMSY 10
Db      188 IVPLGLSY 195
      ||| |||
      TYPE: PRT
      ORGANISM: Homo sapiens
      FEATURE:
      NAME/KEY: SITE
      LOCATION: (243)
      OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
      NAME/KEY: SITE
      LOCATION: (277)
      OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

US-09-925-731-2
; Sequence 2, Application US/09925731
; Patent No. US20020090622A1
; GENERAL INFORMATION:
; APPLICANT: ADEKUN, ANTHONI MONISOLA
; APPLICANT: AMROSE, HELEN JEAN
; APPLICANT: CRESSWELL, CARL JOHN
; APPLICANT: DUDLEY, ADAM JESTON
; TITLE OF INVENTION: CHEMICAL COMPOUNDS
; FILE REFERENCE: DJB/009901/0282795
; CURRENT APPLICATION NUMBER: US/09/925,731
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/226,909
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 691
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-731-2

Query Match      61.5%; Score 32; DB 10; Length 691;
Best Local Similarity 62.5%; Pred. No. 3.2e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      3 VVPXGMSY 10
Db      188 IVPLGLSY 195
      ||| |||
      TYPE: PRT
      ORGANISM: Homo sapiens
      FEATURE:
      NAME/KEY: SITE
      LOCATION: (243)
      OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
      NAME/KEY: SITE
      LOCATION: (277)
      OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

US-09-815-242-10384
; Sequence 10384, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625

```

us-09-909-164-10.rapb

Wed Jun 11 15:41:55 2003

```

; NAME/KEY: misc_feature
; LOCATION: (7)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-092-154-878

Query Match      59.6%; Score 31; DB 9; Length 53;
Best Local Similarity 66.7%; Pred. No. 33;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      3 VVPXGMSYS 11
      ||| :|||
Db      39 VVPTAVSYS 47

RESULT 24
US-09-764-847-878
; Sequence 878, Application US/09764847
; Patent No. US20020132767A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC009
; CURRENT APPLICATION NUMBER: US/09/764,847
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2003
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 878
; LENGTH: 53
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (7)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-847-878

Query Match      59.6%; Score 31; DB 10; Length 53;
Best Local Similarity 66.7%; Pred. No. 33;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      3 VVPXGMSYS 11
      ||| :|||
Db      39 VVPTAVSYS 47

RESULT 25
US-09-948-080-14
; Sequence 14, Application US/09948080
; Patent No. US20020102702A1
; GENERAL INFORMATION:
; APPLICANT: VAN DER OSTEN, CLAUS
; APPLICANT: HALKIER, TORDEN
; APPLICANT: ANDERSEN, CARSTEN
; APPLICANT: BAUDITZ, PETER
; APPLICANT: HANSEN, PETER KAMP
; TITLE OF INVENTION: PROTEASE VARIANTS AND COMPOSITIONS
; FILE REFERENCE: 4946,200-US
; CURRENT APPLICATION NUMBER: US/09/948,080
; CURRENT FILING DATE: 2001-09-06
; PRIOR APPLICATION NUMBER: US/08/963,851
; PRIOR FILING DATE: 1997-11-04
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 14
; LENGTH: 59
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-948-080-14

Query Match      59.6%; Score 31; DB 10; Length 59;
Best Local Similarity 45.5%; Pred. No. 37;
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10384
; LENGTH: 1377
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-815-242-10384

Query Match      61.5%; Score 32; DB 10; Length 1377;
Best Local Similarity 45.5%; Pred. No. 6.7e+02;
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY      1 EEWVPMGMSYS 11
      |::| :|||
Db      369 EQNPAGLSYT 379

RESULT 22
US-10-151-736-4
; Sequence 4, Application US/10151736
; Publication No. US20020192160A1
; GENERAL INFORMATION:
; APPLICANT: Callaghan, Michelle J.
; APPLICANT: Sutherland, Lindfield
; APPLICANT: Watts, Colin K.
; TITLE OF INVENTION: No. US20020192160alel Human Tumour Suppressor Gene
; FILE REFERENCE: RICE-010CON
; CURRENT APPLICATION NUMBER: US/10/151,736
; CURRENT FILING DATE: 2002-05-15
; PRIOR APPLICATION NUMBER: 09/403,402
; PRIOR FILING DATE: 1999-10-21
; PRIOR APPLICATION NUMBER: PCT/AU98/00280
; PRIOR FILING DATE: 1998-04-20
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 2799
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-151-736-4

Query Match      61.5%; Score 32; DB 9; Length 2799;
Best Local Similarity 60.0%; Pred. No. 1.4e+03;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      2 EEWVPMGMSYS 11
      ||| :|||
Db      2096 EVLPKMSYA 2105

RESULT 23
US-10-092-154-878
; Sequence 878, Application US/10092154
; Publication No. US20030054375A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC009C1
; CURRENT APPLICATION NUMBER: US/10/092,154
; CURRENT FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 2003
; Prior application removed - See File Wrapper or Palm
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 878
; LENGTH: 53
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:

```

Wed Jun 11 15:41:55 2003

us-09-909-164-10.rapb

Page 8

QY 1 EEVVPXGMSYS 11  
| : | | : | |  
Db 38 EKHIPGGLEYS 48

Search completed: June 10, 2003, 14:35:43  
Job time : 16.0714 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2003, 13:31:15 ; Search time 11.2143 Seconds  
(without alignments)  
94.297 Million cell updates/sec

Title: US-09-909-164-10  
Perfect score: 52  
Sequence: 1 EEVVPXGMSYS 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues  
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 75 summaries

Database : PIR\_73:  
1: Piri:  
2: Piri2:  
3: Piri3:  
4: Piri4:

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	38	73.1	3472	T31308	hypothetical 367k
2	37	71.2	840	T39116	probable sulfate p
3	37	71.2	877	T40413	sulfate permease -
4	36	69.2	102	A42452	VI protein - tobac
5	36	69.2	1498	B97355	DNA segregation Ar
6	35	67.3	225	S57810	hypothetical prote
7	35	67.3	425	T24111	hypothetical prote
8	35	67.3	670	S22293	zinc finger protei
9	35	67.3	749	H82691	topoisomerase IV s
10	35	67.3	2717	A34203	DNA-binding protei
11	34	65.4	156	S54619	hypothetical prote
12	34	65.4	252	H69491	cell division inhi
13	34	65.4	544	C82900	probable ABC subst
14	33	63.5	94	I40758	hypothetical prote
15	33	63.5	116	E90544	50S ribosomal prot
16	33	63.5	165	D69493	hypothetical prote
17	33	63.5	253	C81374	hypothetical prote
18	33	63.5	259	T34536	hypothetical prote
19	33	63.5	284	S75817	hypothetical prote
20	33	63.5	298	T47670	beta-ketoacyl-ACP
21	33	63.5	368	F72281	hypothetical prote
22	33	63.5	426	D82163	3-phosphoshikimate
23	33	63.5	466	T43653	cdc37 protein - fi
24	33	63.5	653	T48352	iron(III) ABC tran
25	33	63.5	890	A30481	bacteriocin BCN5 -
26	33	63.5	1028	A33286	ATP-dependent DNA
27	33	63.5	1152	D87046	conserved hypothet
28	33	63.5	1394	A35626	transforming growt
29	33	63.5	1401	G82336	DNA-directed RNA p

30	63.5	1548	2	T04456
31	63.5	1712	2	A38261
32	61.5	84	2	E97333
33	61.5	175	2	PQ0616
34	61.5	223	2	T01457
35	61.5	279	2	B72481
36	61.5	288	2	JC4011
37	61.5	288	2	I58372
38	61.5	289	2	A41984
39	61.5	289	2	A42822
40	61.5	291	2	S57925
41	61.5	291	2	JC4579
42	61.5	291	2	S62730
43	61.5	292	2	B42822
44	61.5	292	2	A38977
45	61.5	295	2	A35523
46	61.5	295	2	JC2342
47	61.5	347	2	I51120
48	61.5	363	2	D69551
49	61.5	427	2	F64064
50	61.5	449	2	A9286
51	61.5	498	2	B90604
52	61.5	525	2	D98311
53	61.5	525	2	AF2971
54	61.5	726	2	T44000
55	61.5	726	2	T44187
56	61.5	759	2	S25330
57	61.5	889	2	S22659
58	61.5	922	2	AG1827
59	61.5	993	1	P1VXTA
60	61.5	1306	2	S22624
61	61.5	1377	2	C65159
62	61.5	1377	2	E86034
63	61.5	1394	2	H92336
64	61.5	1397	2	A85570
65	61.5	1397	2	C64805
66	61.5	1399	2	A99720
67	61.5	1409	2	F91187
68	61.5	1411	2	E65145
69	61.5	1411	2	S00485
70	61.5	1482	1	VKLJ51
71	59.6	124	1	VKLJ51
72	59.6	133	2	A71173
73	59.6	208	1	D70764
74	59.6	224	2	G70709
75	59.6	224	2	F87186

ALIGNMENTS

RESULT 1

T31308

hypothetical 367k protein - Cenarchaeum symbiosum

C:Species: Cenarchaeum symbiosum

C:Date: 11-Jan-2000 #sequence\_revision 11-Jan-2000 #text\_change 18-Feb-2000

C:Accession: T31308

R:Schleper, C.; DeLong, E.F.; Preston, C.M.; Feldman, R.A.; Wu, K.Y.; Swanson, R.V.

J. Bacteriol. 180, 5003-5009, 1998

A:Title: Genomic analysis reveals chromosomal variation in natural populations of t

A:Reference number: 220994; PMID:98422450; PMID:9748430

A:Accession: T31308

A:Status: preliminary; translated from GB/EMBL/DBD

A:Molecule type: DNA

A:Residues: 1-3472 <SCH>

A:Cross-references: EMBL:AF083072; NID:g3599393; PID:g3599394; PIDN:AACG2699.1

C:Superfamily: Cenarchaeum symbiosum hypothetical 367k protein

Query Match 73.1%; Score 38; DB 2; Length 3472;  
Best Local Similarity 54.5%; Pred. No. 59;  
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

1 EEVVPXGMSYS 11

Db 2294 EDVIPRGISFS 2304

1:1:1 1:1:1

RESULT 2

T39116

probable sulfate permease - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe

C>Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Dec-1999

C:Accession: T39116

R:Hunt, C.; Aves, S.; McDougall, R.C.; Rajandream, M.A.; Barrell, B.G.

submitted to the EMBL Data Library, November 1999

A:Reference number: 221829

A:Accession: T39116

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-840 <HUN>

A:Cross-references: EMBL:AL132779; PIDN:CAB60015.1; GSPDB:GN000066; SPDB:SPAC869.05c

A:Experimental source: strain 972h-; cosmid c869

C:Genetics:

A:Gene: SPDB:SPAC869.05c

A:Map position: 1

Query Match 71.2%; Score 37; DB 2; Length 840;

Best Local Similarity 77.8%; Pred. No. 21;

Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMSYS 11

111 1111

Db 135 VVPQGMSYA 143

RESULT 3

T40413

sulfate permease - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe

C>Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Dec-1999

C:Accession: T40413

R:Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Jimenez Martinez, J.

submitted to the EMBL Data Library, August 1998

A:Reference number: 221926

A:Accession: T40413

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-877 <LYN>

A:Cross-references: EMBL:AL031261; PIDN:CAA20298.1; GSPDB:GN000067; SPDB:SPBC3H7.02

A:Experimental source: strain 972h-; cosmid c3H7

C:Genetics:

A:Gene: SPDB:SPBC3H7.02

A:Map position: 2

Query Match 71.2%; Score 37; DB 2; Length 877;

Best Local Similarity 77.8%; Pred. No. 22;

Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMSYS 11

111 1111

Db 148 VVPQGMSYA 156

RESULT 4

A42452

V1 protein - tobacco yellow dwarf virus (strain Australia)

C:Species: tobacco yellow dwarf virus

C>Date: 15-Jan-1993 #sequence\_revision 15-Jan-1993 #text\_change 08-Oct-1999

C:Accession: A42452

R:Morris, B.A.M.; Richardson, K.A.; Haley, A.; Zhan, X.; Thomas, J.E.

Virology 187, 633-642, 1992

A:Title: The nucleotide sequence of the infectious cloned DNA component of tobacco yellow dwarf virus

A:Reference number: A42452; MUID:92188538; PMID:1546458

A:Accession: A42452

A:Molecule type: DNA

A:Residues: 1-102 <MOR>

A:Cross-references: GB:M81103; NID:g335283; PIDN:AAA47947.1; PID:g335284

Query Match 69.2%; Score 36; DB 2; Length 102;

Best Local Similarity 60.0%; Pred. No. 3.5;

Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 EVVPXGMSYS 11

1111 1:11

Db 7 QVVPNGINYS 16

RESULT 5

B97355

DNA segregation ATPase, FtsK/SpoIIIE family, YUKA B. subtilis ortholog [imported]

C:Species: Clostridium acetobutylicum

C>Date: 14-Sep-2001 #sequence\_revision 14-Sep-2001 #text\_change 14-Sep-2001

C:Accession: B97355

R:Nolling, J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.

J. Bacteriol. 183, 4823-4838, 2001

A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clostridium acetobutylicum

A:Reference number: A96900; MUID:21359325; PMID:21359325

A:Accession: B97355

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1498 <KUR>

A:Cross-references: GB:AE001437; PIDN:AAK81629.1; PID:g15026814; GSPDB:GN000168

A:Experimental source: Clostridium acetobutylicum ATCC824

C:Genetics:

A:Gene: CAC3709

Query Match 69.2%; Score 36; DB 2; Length 1498;

Best Local Similarity 60.0%; Pred. No. 63;

Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10

1: 1 1111

Db 1276 EOKIPMGMSY 1285

RESULT 6

S57810

hypothetical protein precursor (clone TPPI1) - tomato

C:Species: Lycopersicon esculentum (tomato)

C>Date: 28-Oct-1995 #sequence\_revision 03-Nov-1995 #text\_change 21-Jul-2000

C:Accession: S57810

R:Milligan, S.B.; Gasser, C.S.

Plant Mol. Biol. 28, 691-711, 1995

A:Title: Nature and regulation of pistil-expressed genes in tomato.

A:Reference number: S57808; MUID:95375233; PMID:7647301

A:Accession: S57810

A>Status: preliminary; nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-225 <MIL>

A:Cross-references: EMBL:U20592; NID:g924625; PIDN:AAA80497.1; PID:g924626

C:Superfamily: plant Kunitz-type proteinase inhibitor

Query Match 67.3%; Score 35; DB 2; Length 225;

Best Local Similarity 54.5%; Pred. No. 13;

Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMSYS 11

1111 1:1

Db 32 DEVVPNGKTYA 42

RESULT 7

T24111

hypothetical protein R10D12.10 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999

C:Accession: T24111

R:Percy, C.

submitted to the EMBL Data Library, October 1996

A:Reference number: Z19842  
A:Accession: T24111  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-425 <WIL>  
A:Cross-references: EMBL:Z81109; PIDN:CAB03241.1; GSPDB:GN00023; CESP:R10D12.10  
A:Experimental source: Clone R10D12  
C:Genetics:  
A:Gene: CESP:R10D12.10  
A:Map position: 5  
A:Introns: 23/3; 56/3; 113/3; 257/2

Query Match 67.3%; Score 35; DB 2; Length 425;

Best Local Similarity 50.0%; Pred. No. 26;

Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EVVFXGMSY 10

DB 335 EQIVPGGLGY 344

RESULT 8

S22293  
Zinc finger protein AT-BP2 - rat (fragment)  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 29-Jan-1998 #sequence\_revision 06-Feb-1998 #text\_change 20-Sep-1999  
C:Accession: S22293; J78656  
R:Mitchellmore, C.; Traboni, C.; Cortese, R.  
Nucleic Acids Res. 19, 141-147, 1991  
A:Title: Isolation of two cDNAs encoding zinc finger proteins which bind to the alpha 1-  
A:Reference number: I58280; MUID:91187610; PMID:1901405  
A:Accession: S22293  
A>Status: nucleic acid sequence not shown  
A:Molecule type: mRNA

A:Residues: 1-670 <MIT>  
A:Cross-references: EMBL:X54250; NID:957519; PIDN:CAA38151.1; PID:957520

A>Note: The authors did not translate the codon for residue 1

C:Superfamily: HIV-EP2 enhancer-binding protein

C:Keywords: DNA binding; transcription regulation; zinc finger

Query Match 67.3%; Score 35; DB 2; Length 670;

Best Local Similarity 66.7%; Pred. No. 43;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMSYS 11

DB 376 VVPAGLTYS 384

RESULT 9

H82691  
topoisomerase IV subunit XF1353 [Imported] - Xylella fastidiosa (strain 9a5c)  
C:Species: Xylella fastidiosa  
C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Sep-2000  
C:Accession: H82691  
R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen  
Nature 406, 151-157, 2000  
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.  
A:Reference number: A82515; MUID:20365717; PMID:10910347  
A>Note: for a complete list of authors see reference number A59328 below  
A:Accession: H82691  
A>Status: preliminary  
A:Molecule type: DNA

A:Residues: 1-749 <SIM>  
A:Cross-references: GB:AE003967; GB:AE003849; NID:99106347; PIDN:AAF84162.1; GSPDB:GN001

A:Experimental source: strain 9a5c

R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A

Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H

as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.

submitted to GenBank, June 2000

A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm

J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig

chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martin  
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C  
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, S  
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sa  
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Sil  
M.; Tsubako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.I  
A:Reference number: A59328  
A:Contents: annotation  
C:Genetics:  
A:Gene: XF1353  
C:Superfamily: DNA topoisomerase (ATP-hydrolyzing) chain A; phage T4 DNA topoisom

Query Match 67.3%; Score 35; DB 2; Length 749;

Best Local Similarity 77.8%; Pred. No. 48;

Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 EVVFXGMSY 10

DB 526 EVDPSGMSY 534

RESULT 10

A34203  
DNA-binding protein PRDII-BF1 - human  
N:Alternate names: major histocompatibility complex enhancer-binding protein 1  
C:Species: Homo sapiens (man)  
C:Date: 22-Jun-1990 #sequence\_revision 22-Jun-1990 #text\_change 20-Sep-1999  
C:Accession: A34203; A34779  
R:Fan, C.M.; Maniatis, T.  
Genes Dev. 4, 29-42, 1990  
A:Title: A DNA-binding protein containing two widely separated zinc finger motifs  
A:Reference number: A34203; MUID:90169514; PMID:2106471  
A:Accession: A34203  
A>Status: preliminary  
A:Molecule type: mRNA

A:Residues: 1-2717 <FAN>  
A:Cross-references: EMBL:X51435; NID:938017; PIDN:CAA35798.1; PID:938018

R:Balwin Jr., A.S.; LeClair, K.P.; Singh, H.; Sharp, P.A.

Mol. Cell. Biol. 10, 1406-1414, 1990

A:Title: A large protein containing zinc finger domains binds to related sequence

A:Reference number: A34779; MUID:90205817; PMID:2108316

A:Accession: A34779

A>Status: preliminary; nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 801-1072; N', 1074-1168, 'K', 1170-1225, 'V', 1227-1434, 'N', 1436-1607, 'I', 1

A:Cross-references: GB:M32019

A:Superfamily: HIV-EP2 enhancer-binding protein

C:Keywords: DNA binding; transcription regulation; zinc finger

Query Match 67.3%; Score 35; DB 2; Length 2717;

Best Local Similarity 66.7%; Pred. No. 1.9e+02;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMSYS 11

DB 2405 VVPAGLTYS 2413

RESULT 11

S54619  
hypothetical protein YOR013w - yeast (Saccharomyces cerevisiae)  
N:Alternate names: hypothetical protein O2612; hypothetical protein YOL303.3  
C:Species: Saccharomyces cerevisiae  
C:Date: 08-Jul-1995 #sequence\_revision 01-Sep-1995 #text\_change 19-Apr-2002  
C:Accession: S54619; S66879  
R:de Haan, M.; Maarse, A.C.; Grivell, L.A.  
submitted to the EMBL Data Library, May 1995  
A:Reference number: S54619  
A:Accession: S54619  
A:Molecule type: DNA  
A:Residues: 1-156 <DEH>  
A:Cross-references: EMBL:X87331; NID:91041652; PIDN:CAA60762.1; PID:g829123  
R:de Haan, M.; Grivell, L.A.; Maarse, A.C.

submitted to the Protein Sequence Database, July 1996  
 A:Reference number: S66877  
 A:Accession: S66879  
 A:Molecule type: DNA  
 A:Residues: 1-156 <DEW>  
 A:Cross-references: EMBL:Z74920; NID:g1420109; PIDN:CAA99201.1; PID:g1420111; MIPS:YOR01  
 A:Experimental source: strain S288C  
 C:Genetics:  
 A:Cross-references: SGD:S0005539  
 A:Map position: 15R  
 C:Superfamily: hypothetical protein YOR013w

Query Match 65.4%; Score 34; DB 2; Length 156;  
 Best Local Similarity 66.7%; Pred. No. 14;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 EVVPXGMSY 10  
 ||| |||  
 Db 50 EVMLPGMDY 58

## RESULT 12

H69491  
 cell division inhibitor (mind-2) homolog - Archaeoglobus fulgidus  
 C:Species: Archaeoglobus fulgidus  
 C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 26-Aug-1999  
 C:Accession: H69491  
 R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson  
 ; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.  
 ; Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.  
 Nature 390, 364-370, 1997  
 A:Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.  
 Smith, H.O.; Woese, C.R.; Venter, J.C.  
 A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo  
 A:Reference number: A69250; MUID:98049343; PMID:9389475  
 A:Accession: H69491  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-252 <KLE>  
 A:Cross-references: GB:AE000970; GB:AE000782; NID:g2689293; PIDN:AA889318.1; PID:g264860  
 C:Superfamily: cell division inhibitor mind

Query Match 55.4%; Score 34; DB 2; Length 252;  
 Best Local Similarity 75.0%; Pred. No. 24;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 EVVPXGMS 9  
 ||| |||  
 Db 81 EVIPAGMS 88

## RESULT 13

C82900  
 probable ABC substrate-binding protein, iron U0359 [imported] - Ureaplasma urealyticum  
 C:Species: Ureaplasma urealyticum  
 C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Sep-2000  
 C:Accession: C82900  
 R:Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Cassell, G.H.  
 Submitted to GenBank, February 2000  
 A:Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a min  
 A:Reference number: A82870  
 A:Accession: C82900  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-344 <GLA>  
 A:Cross-references: GB:AE002133; GB:AF222894; NID:g6899339; PIDN:AAF30768.1; GSPDB:GN001  
 A:Experimental source: serovar 3; biovar 1  
 C:Genetics:  
 A:Gene: ABCsbp-5; U0359  
 A:Genetic code: SGC3

Query Match 65.4%; Score 34; DB 2; Length 544;  
 Best Local Similarity 70.0%; Pred. No. 55;

Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 EEVVPXGMSY 10  
 ||||| |||  
 Db 135 EEVVPHYLSY 144

## RESULT 14

I40758  
 hypothetical protein 1 - Campylobacter jejuni (fragment)  
 C:Species: Campylobacter jejuni  
 C:Date: 16-Aug-1996 #sequence\_revision 16-Aug-1996 #text\_change 08-Oct-1999  
 C:Accession: I40758; S47317  
 R:Hani, E.K.; Chan, V.L.  
 J. Bacteriol. 177, 2396-2402, 1995  
 A:Title: Expression and characterization of Campylobacter jejuni benzoylglycine am  
 A:Reference number: I40758; MUID:95247673; PMID:7730270  
 A:Accession: I40758  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-94 <RES>  
 A:Cross-references: EMBL:Z36940; NID:g535805; PIDN:CAA85392.1; PID:g535806

Query Match 63.5%; Score 33; DB 2; Length 94;  
 Best Local Similarity 55.6%; Pred. No. 13;  
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 EVVPXGMSY 10  
 :| ||||  
 Db 26 DIFPMSY 34

## RESULT 15

E90544  
 50S ribosomal protein L20 [imported] - Mycoplasma pulmonis (strain UAB CTIP)  
 C:Species: Mycoplasma pulmonis  
 C:Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 03-Aug-2001  
 C:Accession: E90544  
 R:Chambaud, I.; Heilig, R.; Ferris, S.; Barbe, V.; Sanson, D.; Galisson, F.; Moszer,  
 Nucleic Acids Res. 29, 2145-2153, 2001  
 A:Title: The complete genome sequence of the murine respiratory pathogen Mycoplasma  
 A:Reference number: A99512; MUID:21267165; PMID:11353084  
 A:Accession: E90544  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-116 <KUR>  
 A:Cross-references: GB:AL445566; PID:g14089674; PIDN:CAC13434.1; GSPDB:GN00153  
 A:Experimental source: strain UAB CTIP  
 C:Genetics:  
 A:Gene: MYPUL\_2610  
 A:Genetic code: SGC3  
 C:Superfamily: Escherichia coli ribosomal protein L20

Query Match 63.5%; Score 33; DB 2; Length 116;  
 Best Local Similarity 77.8%; Pred. No. 17;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 VVPXGMSYS 11  
 | | |||||  
 Db 68 VRPLGMSYS 76

## RESULT 16

D69493  
 hypothetical protein AF1949 - Archaeoglobus fulgidus  
 C:Species: Archaeoglobus fulgidus  
 C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 04-Mar-2000  
 C:Accession: D69493  
 R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Do  
 ; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.  
 ; Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.  
 Nature 390, 364-370, 1997  
 A:Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes,

Smith, H.O.; Woese, C.R.; Venter, J.C.  
A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaea  
A:Reference number: A69250; MUID:98049343; PMID:9389475  
A:Accession: D69493  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-165 <KLE>  
A:Cross-references: GB:AE000968; GB:AE000782; NID:92689291; PIDN:AAB89307.1; PID:g264859  
C:Superfamily: Archaeoglobus fulgidus hypothetical protein AFI949

Query Match 63.5%; Score 33; DB 2; Length 165;  
Best Local Similarity 60.0%; Pred. No. 25;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EVVPXGMSY 10  
||| ||| |||  
DB 60 EESIPDGASY 69

RESULT 17  
CB1374  
hypothetical protein Cj0990c [imported] - Campylobacter jejuni (strain NCTC 11168)  
C:Species: Campylobacter jejuni  
C:Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 03-Jun-2002  
C:Accession: CB1374  
R:Parkhill, J.; Wren, B.W.; Mungall, K.; Kettle, J.M.; Churcher, C.; Basham, D.; Chilling  
C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVleet, A.; Whitehead, S.; Barrell  
Nature 403, 665-668, 2000  
A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hyp  
A:Reference number: AB1250; MUID:20150912; PMID:10688204  
A:Accession: CB1374  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-253 <PAR>  
A:Cross-references: GB:AL111168; NID:g6968128; PIDN:CAB73246.1; PID:g696842  
A:Experimental source: serotype O2, strain NCTC 11168  
C:Genetics:  
A:Gene: Cj0990c

Query Match 63.5%; Score 33; DB 2; Length 253;  
Best Local Similarity 55.6%; Pred. No. 39;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 EVVPXGMSY 10  
:: | |||  
DB 185 DIPPGMSY 193

RESULT 18  
T34536  
hypothetical protein DKFZp434C031.1 - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 29-Oct-1999  
C:Accession: T34536  
R:Poustka, A.; Wellenreuther, R.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.  
submitted to the Protein Sequence Database, October 1999  
A:Reference number: Z21540  
A:Accession: T34536  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-259 <POU>  
A:Cross-references: EMBL:AL122063  
A:Experimental source: adult testis; clone DKFZp434C031  
C:Genetics:  
A:Note: DKFZp434C031.1

Query Match 63.5%; Score 33; DB 2; Length 259;  
Best Local Similarity 60.0%; Pred. No. 40;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 EVVPXGMSY 11  
||| ||| |||  
DB 22 EVAPAGASY 31

RESULT 19  
S75817  
hypothetical protein slr1275 - Synechocystis sp. (strain PCC 6803)  
C:Species: Synechocystis sp.  
A:Variety: PCC 6803  
C:Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 08-Oct-1999  
C:Accession: S75817  
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asanizu, E.; Nakamura, Y.; Miyajima,  
O.; K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Ya  
DNA Res. 3, 109-136, 1996  
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis  
S.  
A:Reference number: S74322; MUID:97061201; PMID:8905231  
A:Accession: S75817  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-284 <KAN>  
A:Cross-references: EMBL:D90913; GB:AB001339; NID:gl65348; PIDN:BAAL8276.1; PID:d10  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 63.5%; Score 33; DB 2; Length 284;  
Best Local Similarity 55.6%; Pred. No. 44;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMSYS 11  
||| ||| |||  
DB 208 VIPAGVSYT 216

RESULT 20  
T47670  
beta-ketoacyl-ACP reductase-like protein - Arabidopsis thaliana  
N:Alternate names: protein T26112.190  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 18-Aug-2000  
C:Accession: T47670  
R:Monfort, A.; Casacuberta, E.; Puigdomenech, P.; Mewes, H.W.; Lemcke, K.; Mayer, K  
submitted to the Protein Sequence Database, February 2000  
A:Reference number: Z24471  
A:Accession: T47670  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-298 <MON>  
A:Cross-references: EMBL:AL132954  
A:Experimental source: cultivar Columbia; BAC clone T26112  
C:Genetics:  
A:Map position: 3  
A:Introns: 25/3  
A:Note: T26112.190  
C:Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology

Query Match 63.5%; Score 33; DB 2; Length 298;  
Best Local Similarity 55.6%; Pred. No. 47;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMSYS 11  
||| ||| |||  
DB 187 IVPGLAYS 195

RESULT 21  
F72281  
hypothetical protein TM1216 - Thermotoga maritima (strain MSB8)  
C:Species: Thermotoga maritima  
C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 21-Jul-2000  
C:Accession: F72281  
R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; I  
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardso  
C.M.  
Nature 399, 323-329, 1999  
A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from geno

A:Reference number: A72200; MUID:99287316; PMID:10360571

A:Accession: F72281

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-368 <ARN>

A:Cross-references: GB:AE001778; GB:AE000512; NID:g4981757; PIDN:AAD36291.1; PID:g498176

A:Experimental source: strain MS88

C:Genetics:

A:Gene: TM1216

C:Superfamily: NADH dehydrogenase (ubiquinone) 49K protein

Query Match 63.5%; Score 33; DB 2; Length 368;

Best Local Similarity 55.6%; Pred. No. 59;

Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMSYS 11

Db 294 IVPKGMAYA 302

RESULT 22

D82163

3-phosphoshikimate 1-carboxyvinyltransferase VC1732 [imported] - Vibrio cholerae (strain

C:Species: Vibrio cholerae

C>Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Feb-2001

C:Accession: D82163

R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;

chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers, P.

l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000

A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

A:Reference number: A82035; MUID:20406833; PMID:10952301

A:Accession: D82163

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-426 <HEI>

A:Cross-references: GB:AE004251; GB:AE003852; NID:g9656248; PIDN:AAF94882.1; GSPDB:GN001

A:Experimental source: serogroup O1; strain N16961; biotype El Tor

C:Genetics:

A:Gene: VC1732

A:Map position: 1

C:Superfamily: 3-phosphoshikimate 1-carboxyvinyltransferase; 3-phosphoshikimate 1-carboxy

Query Match 63.5%; Score 33; DB 2; Length 426;

Best Local Similarity 60.0%; Pred. No. 69;

Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10

Db 223 EFVIPAGQSY 232

RESULT 23

T43653

cdc37 protein - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe

C>Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 18-Feb-2000

C:Accession: T43653; T40791; T43654

R:Westwood, P.K.; Preston, N.C.; Fantes, P.A.

submitted to the EMBL Data Library, March 1999

A:Description: Schizosaccharomyces pombe cdc37 cDNA.

A:Reference number: Z22602

A:Accession: T43653

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: mRNA

A:Residues: 1-466 <WES>

A:Cross-references: EMBL:AJ132376; PIDN:CAB38757.1

R:Wood, V.; Rajandream, M.A.; Barrell, B.G.; Volckaert, G.

submitted to the EMBL Data Library, May 1999

A:Reference number: Z21875

A:Accession: T40791

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-466 <WOO>

A:Cross-references: EMBL:AL049769; PIDN:CAB42371.2; GSPDB:GN00067

A:Experimental source: strain 972h-; cosmid C986

R:Westwood, P.K.; Preston, N.C.; Fantes, P.A.

submitted to the EMBL Data Library, March 1999

A:Description: Schizosaccharomyces pombe cdc37 gene.

A:Reference number: Z22603

A:Accession: T43654

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-466 <WE2>

A:Cross-references: EMBL:AJ132377; PIDN:CAB38758.1

C:Genetics:

A:Gene: cdc37; SPAC9B6.10

A:Map position: 2

A:Introns: 8/2; 17/2; 21/1

Query Match 63.5%; Score 33; DB 2; Length 466;

Best Local Similarity 50.0%; Pred. No. 76;

Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10

Db 98 DSAIPGMSY 107

RESULT 24

D82352

iron(III) ABC transporter, permease protein VC0203 [imported] - Vibrio cholerae (str

C:Species: Vibrio cholerae

C>Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Feb-2001

C:Accession: D82352

R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.

chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Seller

l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000

A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

A:Reference number: A82035; MUID:20406833; PMID:10952301

A:Accession: D82352

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-653 <HEI>

A:Cross-references: GB:AE004110; GB:AE003852; NID:g9654600; PIDN:AAF93379.1; GSPDB:GN

A:Experimental source: serogroup O1; strain N16961; biotype El Tor

C:Genetics:

A:Gene: VC0203

A:Map position: 1

Query Match 63.5%; Score 33; DB 2; Length 653;

Best Local Similarity 66.7%; Pred. No. 11e+02;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXGMS 9

Db 300 EEVVPSGIT 308

RESULT 25

A30481

bacteriocin BCN5 - Clostridium perfringens plasmid pIP404

C:Species: Clostridium perfringens

C>Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 24-Nov-1999

C:Accession: A30481; S03779

R:Garnier, T.; Cole, S.T.

J. Bacteriol. 168, 1189-1196, 1986

A:Title: Characterization of a bacteriocinogenic plasmid from Clostridium perfringens

A:Reference number: JT0354; MUID:87057020; PMID:2877971

A:Accession: A30481

A:Molecule type: DNA

A:Residues: 1-890 <GAR>

A:Cross-references: GB:M32882; GB:J03309; NID:g150738; PIDN:AAA98249.1; PID:g150739

C:Genetics:

A:Gene: bcn

us-09-909-164-10.rpr

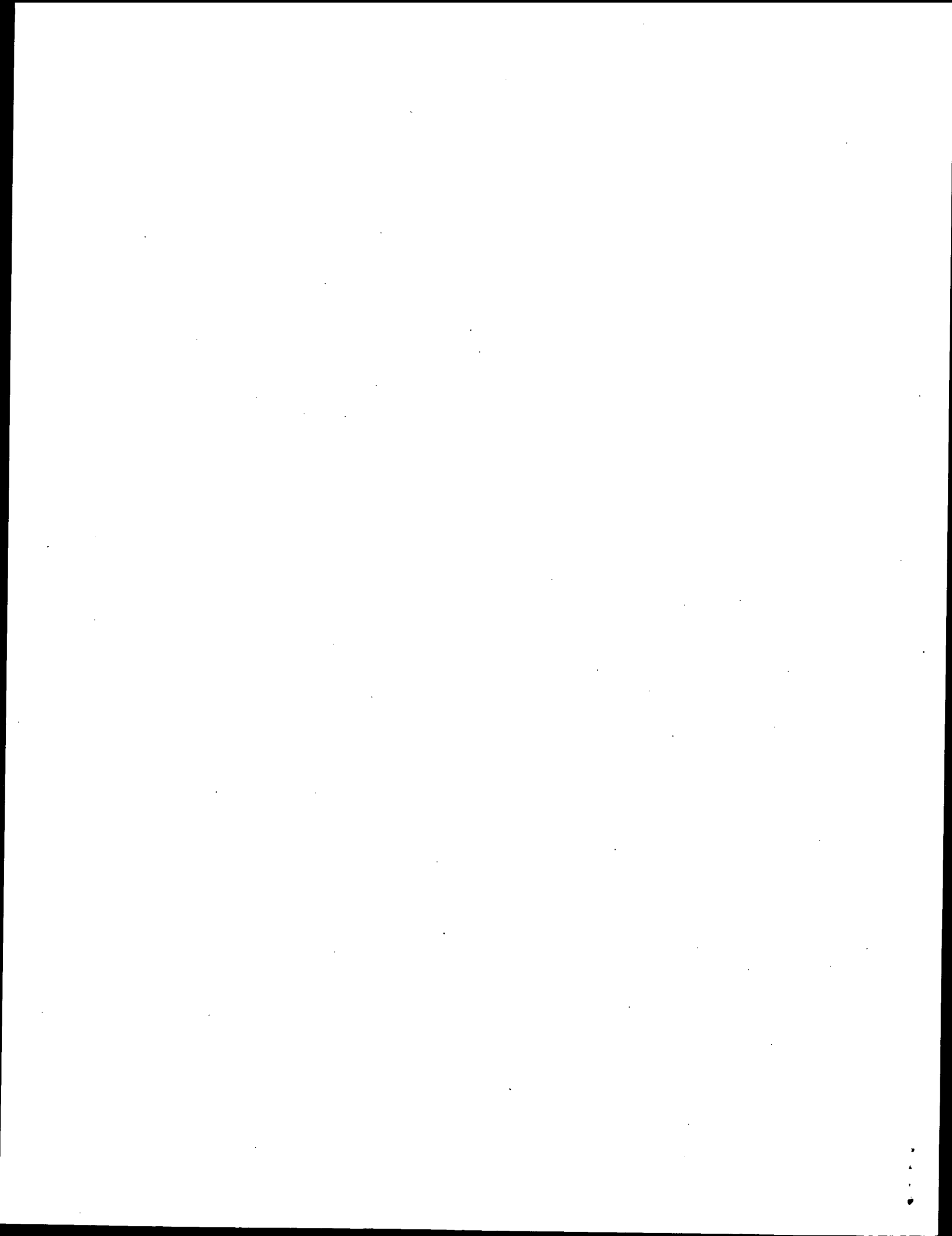
Wed Jun 11 15:42:02 2003

A:Genome: plasmid  
C:Superfamily: Clostridium perfringens plasmid pIP404 bacteriocin BCN5  
C:Keywords: bacteriocin

Query Match 63.5%; Score 33; DB 2; Length 890;  
Best Local Similarity 66.7%; Pred. No. 1.5e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 EVVFXGMSY 10  
|||||!  
Db 170 EVVPGGFTY 178

Search completed: June 10, 2003, 13:49:13  
Job time : 11.2143 secs





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2003, 13:25:04 ; Search time 4.5 Seconds  
(without alignments)  
101.387 Million cell updates/sec

Title: US-09-909-164-10  
Perfect score: 52  
Sequence: 1 EGVVXGMSYS 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 75 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	38	73.1	1058	1 CARB_FUSNN	Q8986 fusobacteri
2	37	71.2	877	1 SULH_SCHPO	074377 schizosacch
3	36	69.2	102	1 Y11K_TYDVA	P31619 tobacco yel
4	36	69.2	1498	1 Y1A9_CLOAB	Q04351 clostridium
5	35	67.3	2717	1 ZEPL_HUMAN	P15822 homo sapien
6	34	65.4	788	1 CY14_NEUCR	P23822 neurospora
7	34	65.4	1499	1 A10C_HUMAN	O60312 homo sapien
8	33	63.5	116	1 RL20_MYCPU	Q98qvo mycoplasma
9	33	63.5	253	1 Y990_CAMJE	P45489 campylobact
10	33	63.5	426	1 AROA_VIBCH	Q9krt0 vibrio chol
11	33	63.5	466	1 GSR2_SCHPO	Q94740 schizosacch
12	33	63.5	478	1 BCN5_CLOPE	Q9nzm5 homo sapien
13	33	63.5	890	1 LTBS_HUMAN	P22084 clostridium
14	33	63.5	1394	1 RPOC_VIBCH	Q9kv29 vibrio chol
15	33	63.5	1401	1 RPOC_VIBCH	Q14766 homo sapien
16	33	63.5	1595	1 LTBL_HUMAN	Q00918 rattus norv
17	33	63.5	1712	1 LTBL_HUMAN	Q04827 rattus norv
18	32	61.5	288	1 CGD2_RAT	P30279 homo sapien
19	32	61.5	289	1 CGD2_HUMAN	P30280 mus musculu
20	32	61.5	289	1 CGD2_MOUSE	Q90459 brachydanio
21	32	61.5	291	1 CGD1_BRARE	P50755 xenopus lae
22	32	61.5	291	1 CGD1_XENLA	P49706 gallus gall
23	32	61.5	291	1 CGD2_CHICK	P53782 xenopus lae
24	32	61.5	291	1 CGD1_XENLA	P55169 gallus gall
25	32	61.5	292	1 CGD1_CHICK	P30281 homo sapien
26	32	61.5	292	1 CGD3_HUMAN	P4385 homo sapien
27	32	61.5	295	1 CGD1_HUMAN	P25322 mus musculu
28	32	61.5	295	1 CGD1_MOUSE	P39948 rattus norv
29	32	61.5	295	1 CGD1_RAT	P44677 haemophilus
30	32	61.5	427	1 TOLB_HAEIN	Q75355 homo sapien
31	32	61.5	529	1 ENP3_HUMAN	O9v616 homo sapien
32	32	61.5	691	1 OAT6_HUMAN	P52384 human herpe
33	32	61.5	726	1 PRTP_HSV60	

34	61.5	759	1	SCT1_YEAST	P2784 saccharomyc
35	61.5	920	1	EDD_RAT	Q62671 rattus norv
36	61.5	993	1	VIA_TAV	P28931 tomato aspe
37	61.5	1377	1	RHSA_ECOLI	P16916 escherichia
38	61.5	1397	1	RHSC_ECOLI	P16918 escherichia
39	61.5	1411	1	RHSE_ECOLI	P16917 escherichia
40	61.5	2799	1	EDD_HUMAN	Q95070 homo sapien
41	61.5	124	1	REV_SIVCZ	P17280 chimpanzee
42	61.5	208	1	COBH_MYCTU	Q10676 mycobacteri
43	61.5	223	1	PURQ_PYRHO	O59619 pyrococcus
44	61.5	224	1	PURQ_HALNI	O9hnu2 halobacteri
45	61.5	224	1	PURQ_MYCLE	O05756 mycobacteri
46	61.5	224	1	PURQ_MYCTU	P71841 mycobacteri
47	61.5	225	1	PURQ_CORAM	Q9rhtx0 corynebacte
48	61.5	240	1	GDIR_ARATH	Q9sf66 arabidopsis
49	61.5	276	1	Y939_METJA	Q58349 methanococc
50	61.5	319	1	YHAL_CRYPA	P10941 cryphonectr
51	61.5	432	1	PURA_YEAST	P80210 saccharomyc
52	61.5	488	1	NOM2_PSEAE	Q9htr0 pseudomonas
53	61.5	670	1	OATP_RAT	Q9uey8 homo sapien
54	61.5	706	1	ADDG_HUMAN	Q9qyb5 mus musculu
55	61.5	706	1	ADDG_MOUSE	Q61586 mus musculu
56	61.5	827	1	PLSB_MOUSE	P71564 rattus norv
57	61.5	828	1	PLSB_RAT	O94489 schizosacch
58	61.5	1047	1	EF3_SCHPO	P00577 escherichia
59	61.5	1407	1	RPOC_ECOLI	P16919 escherichia
60	61.5	1426	1	RHSD_ECOLI	Q9u290 caenorhabdi
61	61.5	2145	1	U520_CAEEL	P19102 xenopus lae
62	61.5	58.7	1	ET2A_XENLA	O28825 archaeoglob
63	61.5	58.7	1	YE47_ARCFU	Q9uhf0 homo sapien
64	61.5	58.7	1	TKNK_HUMAN	O66903 aquifex aeo
65	61.5	58.7	1	ATPE_AQUAE	Q9rgy0 lactobacill
66	61.5	58.7	1	ATPE_LACAC	O26270 methanobact
67	61.5	58.7	1	PURQ_METTH	Q92p11 rhizobium m
68	61.5	58.7	1	PURQ_RHIME	P57489 buchnera ap
69	61.5	58.7	1	RPIA_BUCAI	O59042 methanococc
70	61.5	58.7	1	PURQ_METJA	Q9z1e3 helicobacte
71	61.5	58.7	1	SCOA_HELPJ	P56006 helicobacte
72	61.5	58.7	1	SCOA_HELPJ	P28233 rattus norv
73	61.5	58.7	1	CXA6_RAT	P87033 utillago ma
74	61.5	58.7	1	GBA2_USTMA	P26391 salmonella
75	61.5	58.7	1	RFBB_SALTY	

## ALIGNMENTS

### RESULT 1

ID	CARB_FUSNN	STANDARD	PRT	1058 AA
AC	Q8RG86			
DT	15-JUN-2002 (Rel. 41, Created)			
DT	15-JUN-2002 (Rel. 41, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Carbamoyl-phosphate synthase large chain (EC 6.3.5.5) (Carbamoyl-phosphate synthetase ammonia chain).			
GN	CARB OR F0422.			
OS	Fusobacterium nucleatum (subsp. nucleatum).			
OC	Bacteria; Fusobacteria; Fusobacterium.			
OX	NCBI_TaxID=76856;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=ATCC 25586;			
RX	MEDLINE=21886394; PubMed=11889109;			
RA	Kapatral V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A., Bhattacharya A., Bartman A., Gardner W., Grechkin G., Zhu L., Vasileva O., Chu L., Kogan Y., Chaga O., Goltsman E., Bernal A., Larsen N., D'Souza M., Walunas T., Pusch G., Haselkorn R., Foustain M., Kyripides N., Overbeek R.;			
RA	"Genome sequence and analysis of the oral bacterium Fusobacterium nucleatum strain ATCC 25586."			
RT	J. Bacteriol. 184:2005-2018(2002).			
CC	- - CATALYTIC ACTIVITY: 2 ATP + L-glutamine + CO(2) + H(2)O = 2 ADP +			

RC STRAIN-972;  
 RX MEDLINE=21848401; PubMed=11859360;  
 RA Wood V., Gilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,  
 RA Sgouros J., Peat N., Hayes J., Baker S., Basham D., Bowman S.,  
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,  
 RA Collins M., Connor R., Cronin A., Davis P., Feitwell T., Fraser A.,  
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,  
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,  
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,  
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,  
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitch E.,  
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,  
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,  
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
 RA Woodward J., Volkstadt G., Aert R., Robben J., Grynopre B.,  
 RA Weijens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,  
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,  
 RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,  
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,  
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Laure V., Mottier S.,  
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
 RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,  
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
 RA Cerretti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,  
 RA Shpakovskii G.V., Usevsky D., Bartell B.G., Nurse P.;  
 RX "the genome sequence of *Schizosaccharomyces pombe*.";  
 RL Nature 415:871-880(2002).  
 CC -!- FUNCTION: HIGH AFFINITY UPTAKE OF SULFATE INTO THE CELL (BY  
 CC SIMILARITY).  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
 CC -!- SIMILARITY: BELONGS TO THE SLC36A FAMILY OF TRANSPORTERS.  
 CC -!- SIMILARITY: CONTAINS 1 STAS DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: AL031261; CAA20298.1; -;  
 CC InterPro: IPR002645; STAS;  
 CC InterPro: IPR001902; Sulfate\_transp.  
 CC Pfam: PF00916; Sulfate\_transp.  
 CC Pfam: PF01740; STAS; 1.  
 CC TIGRfam: TIGR00815; sulp; 1.  
 CC PROSITE: PS01301; SLC26A; 1.  
 CC PROSITE: PS0801; STAS; 1.  
 CC Transport; Transmembrane.  
 CC KW TRANSMEM 133 153 POTENTIAL.  
 CC TRANSMEM 161 181 POTENTIAL.  
 CC FT TRANSMEM 196 206 POTENTIAL.  
 CC FT TRANSMEM 221 241 POTENTIAL.  
 CC FT TRANSMEM 243 263 POTENTIAL.  
 CC FT TRANSMEM 292 312 POTENTIAL.  
 CC FT TRANSMEM 329 349 POTENTIAL.  
 CC FT TRANSMEM 384 404 POTENTIAL.  
 CC FT TRANSMEM 424 444 POTENTIAL.  
 CC FT TRANSMEM 461 481 POTENTIAL.  
 CC FT TRANSMEM 484 504 POTENTIAL.  
 CC FT TRANSMEM 518 538 POTENTIAL.  
 CC FT TRANSMEM 543 563 POTENTIAL.  
 CC FT DOMAIN 594 747 STAS.  
 CC SQ SEQUENCE 877 AA; 96373 MW; 56995A8493371E43 CRC64;  
 Query Match 71.2%; Score 37; DB 1; Length 877;  
 Best Local Similarity 77.8%; Pred. No. 8.3;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 3 VVPXGMGYS 11

```

phosphate + L-glutamate + carbamoyl phosphate.
-1- COFACTOR: Binds three manganese ions (By similarity).
-1- PATHWAY: Arginine biosynthesis.
-1- PATHWAY: Pyrimidine biosynthesis.
-1- SUBUNIT: Composed of two chains; the small (or glutamine) chain promotes the hydrolysis of glutamine to ammonia, which is used by the large (or ammonia) chain to synthesize carbamoyl phosphate (By similarity).
-1- SIMILARITY: BELONGS TO THE CARB FAMILY.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
-----
EMBL; AE010554; AAL94625.1; ALT_INIT.
DR InterPro; IPR005483; CPase_L.
DR InterPro; IPR005479; CPase_L_D2.
DR InterPro; IPR005480; CPase_L_D3.
DR InterPro; IPR005481; CPase_L_N.
DR InterPro; IPR004362; MGS_like.
DR Pfam; PF00289; CPsase_L_chain; 2.
DR Pfam; PF02786; CPsase_L_D2; 2.
DR Pfam; PF02787; CPsase_L_D3; 1.
DR Pfam; PF02142; MGS; 1.
DR PRINTS; PR00098; CPsase.
DR PROSITE; PS00866; CPsase_1; 2.
DR PROSITE; PS00867; CPsase_2; 2.
KW Arginine biosynthesis; Pyrimidine biosynthesis; Ligase; Repeat;
ATP-binding; Manganese; Complete proteome.
FT DOMAIN 1 401
FT DOMAIN 402 546
FT DOMAIN 547 929
FT DOMAIN 930 1058
FT REPEAT 1 546
FT REPEAT 547 1058
FT NP_BIND 153 210
FT NP_BIND 302 352
FT METAL 284 284
FT METAL 298 298
FT METAL 300 300
FT METAL 820 820
FT METAL 832 832
FT SEQUENCE 1058 AA; 117451 MW; ED7037AF77C1E39F CRC64;
Query Match 73.1%; Score 38; DB 1; Length 1058;
Best Local Similarity 60.0%; Pred. No. 6.2;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 2 EVVPGXGMSY 11
Db 190 EIVPNGINYS 199
-----
RESULT 2
SULH_SCHPO STANDARD; PRT; 877 AA.
AC 074377;
ID 15-JUL-1999 (Rel. 38, Created)
ID 15-JUL-1999 (Rel. 38, Last sequence update)
ID 15-JUN-2002 (Rel. 41, Last annotation update)
DE Probable sulfate permease C3H7.02.
DE SPC3H7.02.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
XP [1]
SEQUENCE FROM N.A.

```

Db: 148 VVPGMSYA 156

RESULT 3

Y11K TYDVA STANDARD; PRT; 102 AA.

AC P31619;

DT 01-JUL-1993 (Rel. 26, Created)

DT 01-JUL-1993 (Rel. 26, Last sequence update)

DT 01-OCT-1993 (Rel. 27, Last annotation update)

DE Hypothetical 11.2 kDa protein.

GN V1.

OS Tobacco yellow dwarf virus (strain Australia) (TYDV).

OC Viruses; ssDNA viruses; Geminiviridae; Mastrevirus.

OX NCBI\_TaxID=31599;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=92188538; PubMed=1546458;

RA Morris B.A.M., Richardson K.A., Haley A., Zhan X., Thomas J.E.;

RT "The nucleotide sequence of the infectious cloned DNA component of

RT tobacco yellow dwarf virus reveals features of geminiviruses

RT infecting monocotyledonous plants.";

RL Virology 187:633-642(1992).

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

CC EMBL; M81103; AAA47947.1; ..

DR PIR; A42452; A42452.

DR InterPro; IPR002621; Gemini\_mov.

DR Pfam; PF01708; Gemini\_mov; 1.

KW Hypothetical protein.

SQ SEQUENCE 102 AA; 11178 MW; A40ECF1E0AF55B67 CRC64;

Query Match 69.2%; Score 36; DB 1; Length 102;

Best Local Similarity 60.0%; Pred. No. 1.4;

Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 EVVPGMSYS 11

DB :||| |:||

7 QVVPISGINS 16

RESULT 4

Y1A9 CLOAB STANDARD; PRT; 1498 AA.

AC Q04351;

DT 01-FEB-1994 (Rel. 28, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Hypothetical protein CAC3709.

GN CAC3709

OS Clostridium acetobutylicum.

OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;

OC Clostridium.

OX NCBI\_TaxID=1488;

RN [1]

RP SEQUENCE FROM N.A.

RX STRAIN=ATCC 824 / DSM 792 / VKM B-1787;

RX MEDLINE=21359325; PubMed=11466286;

RA Neelling J., Breton G., Omeichenko M.V., Makarova K.S., Zeng Q.,

RA Gibson R.L., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I.,

RA Tatusov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,

RA Bennett G.N., Koonin E.V., Smith D.R.;

RA "Genome sequence and comparative analysis of the solvent-producing

RT bacterium Clostridium acetobutylicum.";

RL J. Bacteriol. 183:4823-4838(2001).

RN [2]

RP SEQUENCE OF 1-108 FROM N.A.

RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;

RX MEDLINE=93273706; PubMed=8501044;

RA Sauer U., Duerre P.;

RT "Sequence and molecular characterization of a DNA region encoding a

RL J. Bacteriol. 175:3394-3400(1993).

CC -|- SIMILARITY: BELONGS TO THE FTSK/SPOILIE FAMILY.

CC -|- CAUTION: Ref.2 sequence differs from that shown due to frameshifts

CC in positions 76 and 106.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

CC EMBL; AE007866; AAK81629.1; ..

DR EMBL; X65276; CA46379.1; ALT\_FRAME.

DR InterPro; IPR002543; FTSK\_SPOILIE.

DR Pfam; PF01580; FTSK\_SPOILIE; 2.

KW Hypothetical protein; ATP-binding; Complete proteome.

FT NP\_BIND 675 682 ATP (POTENTIAL).

SQ SEQUENCE 1498 AA; 168968 MW; FF42037A335A9649 CRC64;

Query Match 69.2%; Score 36; DB 1; Length 1498;

Best Local Similarity 60.0%; Pred. No. 23;

Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10

DB |:| ||||

1276 EOKIPGMSY 1285

RESULT 5

ZEPL\_HUMAN STANDARD; PRT; 2717 AA.

AC P15822;

DT 01-APR-1990 (Rel. 14, Created)

DT 01-APR-1990 (Rel. 14, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Zinc finger protein 40 (Human immunodeficiency virus type I enhancer-

DE binding protein 1) (HIV-Ep1) (Major histocompatibility complex binding

DE protein 1) (MEP-1) (Positive regulatory domain II binding factor 1)

DE (PRDII-BF1).

GN HIVEP1 OR ZNF40.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=90169514; PubMed=2106471;

RA Fan C.M., Maniatis T.;

RT "A DNA-binding protein containing two widely separated zinc finger

RT motifs that recognize the same DNA sequence.";

RL Genes Dev. 4:29-42(1990).

RN [2]

RP STRUCTURE BY NMR OF 2113-2142.

RX MEDLINE=91064333; PubMed=2248949;

RA Omichinski J.G., Clore G.M., Appella E., Sakaguchi K.,

RA Gronenborn A.M.;

RT "High-resolution three-dimensional structure of a single zinc finger

RL from a human enhancer binding protein in solution.";

RN Biochemistry 29:9324-9334(1990).

RN [3]

RP STRUCTURE BY NMR OF 2087-2142.

RX MEDLINE=92232684; PubMed=1567844;

RA Omichinski J.G., Clore G.M., Robien M., Sakaguchi K., Appella E.,

RA Gronenborn A.M.;

RT "High-resolution solution structure of the double Cys2His2 zinc

finger from the human enhancer binding protein MBP-1.";  
 Biochemistry 31:3907-3917(1992).  
 CC -1- FUNCTION: THIS PROTEIN SPECIFICALLY BINDS TO THE DNA SEQUENCE  
 CC 5'-GGGACTTCC-3' WHICH IS FOUND IN THE ENHANCER ELEMENTS OF  
 CC NUMEROUS VIRAL PROMOTERS SUCH AS THOSE OF SV40, CMV, OR HIV1.  
 CC IN ADDITION, RELATED SEQUENCES ARE FOUND IN THE ENHANCER ELEMENTS  
 CC OF A NUMBER OF CELLULAR PROMOTERS, INCLUDING THOSE OF THE CLASS I  
 CC MHC, INTERLEUKIN-2 RECEPTOR, AND INTERFERON-BETA GENES. IT MAY ACT  
 CC IN T-CELL ACTIVATION.  
 CC -1- SUBCELLULAR LOCATION: Nuclear.  
 CC -1- INDUCTION: BY MITOGEN AND PHORBOL ESTER.  
 CC -1- DOMAIN: CONTAINS TWO SETS OF 2 ZINC-FINGERS, WHICH ARE WIDELY  
 CC SEPARATED AND RECOGNIZE THE SAME DNA SEQUENCE. THERE IS A FIFTH  
 CC ZINC-FINGER IN-BETWEEN.  
 CC -1- SIMILARITY: STRONG, TO HIVP2.  
 CC -----  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X51435; CAA35798.1; .  
 DR PIR; A34203; A34203.  
 DR PDB; 32NF; 15-JAN-92.  
 DR PDB; 42NF; 15-JAN-92.  
 DR PDB; 1BBO; 31-OCT-92.  
 DR TRANSFAC; T00497; .  
 DR Genew; HGNC:4920; HIVP1.  
 DR MIM; 194540; .  
 DR InterPro; IPR000822; Znf\_C2H2.  
 DR Pfam; PF00096; zf-C2H2; 5.  
 DR PRINTS; PRO0048; ZINC\_FINGER.  
 DR SMART; SM00355; ZNF\_C2H2; 4.  
 DR PROSITE; PS00028; ZINC\_FINGER\_C2H2\_1; 4.  
 DR PROSITE; PS0157; ZINC\_FINGER\_C2H2\_2; 4.  
 KW Transcription regulation; Zinc-finger; Metal-binding; DNA-binding;  
 KW Nuclear protein; Repeat; 3D-structure.  
 FT DOMAIN 406 456  
 FT ZINC FINGERS.  
 FT ZN\_FING 406 428  
 FT C2H2-TYPE.  
 FT ZN\_FING 434 456  
 FT C2H2-TYPE.  
 FT DOMAIN 803 806  
 FT POLY-SER.  
 FT ZN\_FING 958 981  
 FT C2HC-TYPE (POTENTIAL).  
 FT DOMAIN 2087 2139  
 FT ZINC FINGERS.  
 FT ZN\_FING 2115 2139  
 FT C2H2-TYPE.  
 FT STRAND 2088 2088  
 FT TURN 2090 2092  
 FT STRAND 2095 2095  
 FT STRAND 2099 2099  
 FT HELIX 2109 2109  
 FT TURN 2115 2116  
 FT STRAND 2123 2124  
 FT STRAND 2127 2135  
 FT HELIX 2127 2135  
 SQ SEQUENCE 2717 AA; 297217 MW; D45D3CA951F8A561 CRC64;  
 Query Match 67.3%; Score 35; DB 1; Length 2717;  
 Best Local Similarity 66.7%; Pred. No. 70;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 QY 3 VVPXGMSYS 11  
 ||| |::|||  
 Db 2405 VVPAGLTYS 2413  
 RESULT 6  
 CY14\_NEUCR  
 ID CY14\_NEUCR STANDARD; PRT; 788 AA.  
 AC P23622;  
 DT 01-NOV-1991 (Rel. 20, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)

15-JUN-2002 (Rel. 41, Last annotation update)  
 DT Sulfate permease II.  
 DE GN CYS-14.  
 OS Neurospora crassa.  
 CC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
 CC Sordariales; Sordariaceae; Neurospora.  
 CC NCBI\_TaxID=5141;  
 RX SEQUENCE FROM N.A.  
 RX MEDLINE=91129256; PubMed=1825178;  
 RA Ketter J.S., Jarai G., Fu Y.-H., Marzluf G.A.;  
 RT "Nucleotide sequence, messenger RNA stability, and DNA recognition  
 RT elements of cys-14, the structural gene for sulfate permease II in  
 RT Neurospora crassa.";  
 RL Biochemistry 30:1780-1787(1991).  
 RN [2]  
 RP PROBABLE REVISIONS.  
 RX MEDLINE=94188926; PubMed=8140616;  
 RA Sandal N.N., Marcker K.A.;  
 RT "Similarities between a soybean nodulin, Neurospora crassa sulphate  
 RT permease II and a putative human tumour suppressor.";  
 RL Trends Biochem. Sci. 19:19-19(1994).  
 CC -1- FUNCTION: UPTAKE OF SULFATE INTO THE CELL.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -1- INDUCTION: Highly expressed, but only in cells subject to sulfur  
 CC limitation, and it is turned on by the positive-acting Cys-3  
 CC sulfur regulatory protein.  
 CC -1- MISCELLANEOUS: SULFATE PERMEASE II IS MAINLY FOUND IN MYCELIA.  
 CC -1- SIMILARITY: BELONGS TO THE SLC26A FAMILY OF TRANSPORTERS.  
 CC -----  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M59167; AAA33615.1; ALT\_SEQ.  
 DR PIR; A37956; A37956.  
 DR InterPro; IPR001902; Sulfate\_transp.  
 DR Pfam; PF00916; Sulfate\_transp; 1.  
 DR TIGRFS; TIGR00815; sulp; 1.  
 DR PROSITE; PS01130; SLC26A; 1.  
 KW Transport; Transmembrane; Glycoprotein.  
 FT TRANSMEM 71 91  
 FT POTENTIAL.  
 FT TRANSMEM 103 123  
 FT POTENTIAL.  
 FT TRANSMEM 128 148  
 FT POTENTIAL.  
 FT TRANSMEM 171 191  
 FT POTENTIAL.  
 FT TRANSMEM 193 213  
 FT POTENTIAL.  
 FT TRANSMEM 271 291  
 FT POTENTIAL.  
 FT TRANSMEM 326 346  
 FT POTENTIAL.  
 FT TRANSMEM 363 383  
 FT POTENTIAL.  
 FT TRANSMEM 451 471  
 FT POTENTIAL.  
 FT TRANSMEM 474 494  
 FT POTENTIAL.  
 FT CARBOHYD 23 23  
 FT CARBOHYD 578 578  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 788 AA; 87864 MW; 4FC604B60798CE77 CRC64;  
 Query Match 65.4%; Score 34; DB 1; Length 788;  
 Best Local Similarity 66.7%; Pred. No. 32;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 QY 3 VVPXGMSYS 11  
 ||| |::|||  
 Db 90 VVPQGMAYA 98  
 RESULT 7  
 A10C\_HUMAN  
 ID A10C\_HUMAN STANDARD; PRT; 1499 AA.  
 AC O60312; Q96914;  
 DT 30-MAY-2000 (Rel. 39, Created)

DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Potential phospholipid-transferring ATPase VC (EC 3.6.3.1) (ATPVC)  
 DE (Aminophospholipid translocase VC).  
 GN ATP10C OR ATPVC OR KIAA0566.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE=2125279; PubMed=11326269;  
 RA Meguro M., Kashiwagi A., Mitsuya K., Nakao M., Kondo I., Saitoh S.,  
 RA Oshimura M.;  
 RA "A novel maternally expressed gene, ATP10C, encodes a putative  
 RT aminophospholipid translocase associated with Angelman syndrome.";  
 RT Nat. Genet. 28:19-20(2001).  
 RL [2]  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE=2131319; PubMed=11353404;  
 RA Hering L.B.K., Kim S.-J., Cook E.H. Jr., Ledbetter D.H.;  
 RA "The human aminophospholipid-transferring ATPase gene ATP10C maps  
 RT adjacent to UBE3A and exhibits similar imprinted expression.";  
 RT Am. J. Hum. Genet. 68:1501-1505(2001).  
 RL [3]  
 RN SEQUENCE OF 337-1499 FROM N.A.  
 RP TISSUE=Brain;  
 RC MEDLINE=98290545; PubMed=9628581;  
 RX Nagase T., Ishikawa K.-I., Miyajima N., Tanaka A., Kotani H.,  
 RA Nomura N., Ohara O.;  
 RA "Prediction of the coding sequences of unidentified human genes. IX.  
 RT The complete sequences of 100 new cDNA clones from brain which can  
 RT code for large proteins in vitro.";  
 RT DNA Res. 5:31-39(1998).  
 RL CC  
 CC -1- CATALYTIC ACTIVITY: ATP + H(2)O = ADP + phosphate.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).  
 CC -1- TISSUE SPECIFICITY: Wide expression, with highest levels in  
 CC kidney, followed by lung, brain, prostate, testis, ovary, and  
 CC small intestine.  
 CC -1- DISEASE: Defects in ATP10C may be a cause of Angelman syndrome  
 CC (AS), also known as 'happy puppet syndrome'.  
 CC -1- SIMILARITY: BELONGS TO THE CATION TRANSPORT ATPASES FAMILY (E1-E2  
 CC ATPASES). SUBFAMILY IV.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: AB051358; BAB47392.1; -  
 CC EMBL: AY029504; AAK33100.1; -  
 CC EMBL: AY029487; AAK33100.1; JOINED.  
 CC EMBL: AY029488; AAK33100.1; JOINED.  
 CC EMBL: AY029489; AAK33100.1; JOINED.  
 CC EMBL: AY029490; AAK33100.1; JOINED.  
 CC EMBL: AY029491; AAK33100.1; JOINED.  
 CC EMBL: AY029492; AAK33100.1; JOINED.  
 CC EMBL: AY029493; AAK33100.1; JOINED.  
 CC EMBL: AY029494; AAK33100.1; JOINED.  
 CC EMBL: AY029495; AAK33100.1; JOINED.  
 CC EMBL: AY029496; AAK33100.1; JOINED.  
 CC EMBL: AY029497; AAK33100.1; JOINED.  
 CC EMBL: AY029498; AAK33100.1; JOINED.  
 CC EMBL: AY029499; AAK33100.1; JOINED.  
 CC EMBL: AY029500; AAK33100.1; JOINED.  
 CC EMBL: AY029501; AAK33100.1; JOINED.  
 CC EMBL: AY029502; AAK33100.1; JOINED.  
 CC EMBL: AY029503; AAK33100.1; JOINED.  
 CC EMBL: AB011138; BAA25492.1; -  
 CC Genew; HGNC:13547; ATP10C.

DR MIM: 605855; -  
 DR MIM: 105830; -  
 DR InterPro: IPR001757; ATPase\_E1-E2.  
 DR InterPro: IPR001454; Hlgnaase/hydriase.  
 DR Pfam: PF00702; Hydrolase; 1.  
 DR PRINTS: PR00119; CATAPASE.  
 DR PROSITE: PS00154; ATPASE\_E1-E2; 1.  
 KW Hydrolase; Transmembrane; Phosphorylation; Magnesium; ATP-binding;  
 KW Multigene family.  
 FT DOMAIN 1 86 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 87 106 POTENTIAL.  
 FT DOMAIN 107 110 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 111 128 POTENTIAL.  
 FT DOMAIN 129 309 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 310 332 POTENTIAL.  
 FT DOMAIN 337 362 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 363 384 POTENTIAL.  
 FT DOMAIN 385 1087 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 1088 1108 POTENTIAL.  
 FT DOMAIN 1109 1119 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 1120 1140 POTENTIAL.  
 FT DOMAIN 1141 1170 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 1171 1192 POTENTIAL.  
 FT DOMAIN 1193 1199 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 1200 1222 POTENTIAL.  
 FT DOMAIN 1223 1228 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 1229 1249 POTENTIAL.  
 FT DOMAIN 1250 1267 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 1268 1292 POTENTIAL.  
 FT DOMAIN 1293 1499 CYTOPLASMIC (POTENTIAL).  
 FT MOD\_RES 427 427 PHOSPHORYLATION (BY SIMILARITY).  
 FT METAL 1031 1031 MAGNESIUM (BY SIMILARITY).  
 FT METAL 1035 1035 MAGNESIUM (BY SIMILARITY).  
 FT DOMAIN 467 470 POLY-GLU.  
 FT CONFLICT 368 368 Q -> R (IN REF. 3).  
 SQ SEQUENCE 1499 AA; 167687 MW; D4996A4D0635A68D CRC64;  
 Query Match 65.4%; Score 34; DB 1; Length 1499;  
 Best Local Similarity 72.7%; Pred. No. 62;  
 Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 EEVVPXGMSYS 11  
 Db 469 EEVVPXGMSYS 11  
 RESULT 8  
 RL20\_MYCPU STANDARD; PRT; 116 AA.  
 AC Q98QV0;  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE 50S ribosomal protein L20.  
 GN RPLT OR MYP2L20.  
 OS Mycoplasma pulmonis.  
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.  
 OX NCBI\_TaxID=2107;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=UAB CTIP;  
 RX MEDLINE=21267165; PubMed=11353084;  
 RA Chamberland I., Heilig R., Ferris S., Barbe V., Samson D., Galisson F.,  
 RA Moszer I., Dydvig K., Wroblewski H., Viari A., Rocha E.P.C.,  
 RA Blanchard A.;  
 RT "The complete genome sequence of the murine respiratory pathogen  
 RT Mycoplasma pulmonis.";  
 RL Nucleic Acids Res. 29:2145-2153(2001).  
 CC -1- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 23S RIBOSOMAL RNA AND IS  
 CC NECESSARY TO THE IN VITRO ASSEMBLY PROCESS OF THE 50S RIBOSOMAL  
 CC SUBUNIT; IT IS NOT INVOLVED IN THE PROTEIN SYNTHESIZING FUNCTIONS  
 CC OF THAT SUBUNIT (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE L20P FAMILY OF RIBOSOMAL PROTEINS.

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AL45563; CAC13434.1; ..
DR MyuList; MYPUL2610; ..
DR InterPro; IPR001081; Ribosomal_L20.
DR Pfam; PF00453; Ribosomal_L20; 1.
DR PRINTS; PR00062; RIBOSOMAL_L20.
DR ProDom; PD002389; Ribosomal_L20; 1.
DR TIGRFAMS; TIGR01032; rplT_bact; 1.
DR PROSITE; PS00937; RIBOSOMAL_L20; 1.
KW Ribosomal protein; rRNA-binding; Complete proteome.
SQ SEQUENCE 116 AA; 13565 MW; C59C748901B18F14 CRC64;

Query Match 63.5%; Score 33; DB 1; Length 116;
Best Local Similarity 77.8%; Pred. No. 7;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 VVPXGMSYS 11
DB 68 VRPLGMSYS 76

RESULT 9
Y990_CAMJE STANDARD; PRT; 253 AA.
AC P45489; Q9PNV0;
DT 01-NOV-1995 (Rel. 32, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein Cj0990c.
GN Cj0990c.
OS Campylobacter jejuni.
OC Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group;
CC Campylobacter.
OX NCBI_TaxID=197;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCCT 11168;
RX MEDLINE=20150912; PubMed=10688204;
RA Parkhill J., Wren B.W., Mungall K., Ketley J.M., Churcher C.,
RA Basham D., Chillingworth T., Davies R.M., Feltwell T., Holtroyd S.,
RA Jagels K., Karlyshev A.V., Moule S., Pallen M.J., Penn C.W.,
RA Quail M.A., Rajandream A.M., Rutherford K.M., van Vliet A.H.M.,
RA Whitehead S., Barrrell B.G.;
RT "The genome sequence of the food-borne pathogen Campylobacter jejuni
RT reveals hypervariable sequences.";
RL Nature 403:665-668(2000).
RP SEQUENCE OF 160-253 FROM N.A.
RC STRAIN=ATCC 43431 / TGH 9011;
RX MEDLINE=95247673; PubMed=7730270;
RA Hani E.K., Chan V.L.;
RT "Expression and characterization of Campylobacter jejuni
RT benzoylglycine amidohydrolase (Hippuricase) gene in Escherichia
RT coli.";
RL J. Bacteriol. 177:2396-2402(1995).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AL139076; CAB73246.1; ..

```

```

DR EMBL; Z36940; CAA85392.1; ..
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 253 AA; 29783 MW; F96D3FF3265F8A6A CRC64;

Query Match 63.5%; Score 33; DB 1; Length 253;
Best Local Similarity 55.6%; Pred. No. 16;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 EYVXPXGMSY 10
DB 185 DIFPXSMSY 193

RESULT 10
AROA_VIBCH STANDARD; PRT; 426 AA.
ID AROA_VIBCH
AC Q9KR80;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE 3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) (5-
DE enolpyruvylshikimate-3-phosphate synthase) (EPSP synthase) (EPSPS).
DE AROA OR VC1732.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=El Tor N16961 / Serotype O1;
RX MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Emolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae.";
RL Nature 406:477-483(2000).
CC -I- CATALYTIC ACTIVITY: Phosphoenolpyruvate + 3-phosphoshikimate =
CC phosphate + 5-O-(1-carboxyvinyl)-3-phosphoshikimate.
CC -I- PATHWAY: Aromatic amino acids biosynthesis; shikimate pathway;
CC sixth step.
CC -I- SUBUNIT: Monomer (By similarity).
CC -I- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -I- SIMILARITY: BELONGS TO THE EPSP SYNTHASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE004251; AAF94882.1; ..
DR TIGR; VC1732; ..
DR InterPro; IPR001986; EPSP_synthase.
DR Pfam; PF00275; EPSP_synthase; 1.
DR ProDom; PD001867; EPSP_synthase; 1.
DR PROSITE; PS00104; EPSP_SYNTHASE_1; 1.
DR PROSITE; PS00885; EPSP_SYNTHASE_2; 1.
KW Aromatic amino acid biosynthesis; Transferrase; Complete proteome.
SQ SEQUENCE 426 AA; 46101 MW; 38852D6483BF1C3 CRC64;

Query Match 63.5%; Score 33; DB 1; Length 426;
Best Local Similarity 60.0%; Pred. No. 27;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EYVXPXGMSY 10
DB 223 EFVIPAGQSY 232

```

Query Match 63.5%; Score 33; DB 1; Length 466;  
 Best Local Similarity 50.0%; Pred. No. 30;  
 Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSY 10  
 : : : : :  
 Db 98 DSAIPGMSY 107  
 : : : : :  
 : : : : :

RESULT 12  
 GSR2\_HUMAN STANDARD; PRT; 478 AA.  
 ID GSR2\_HUMAN STANDARD; PRT; 478 AA.  
 AC Q9NZM5; Q9NP11; Q9NF12; Q9BTC6; Q9HAX6;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Glioma tumor suppressor candidate region gene 2 protein (p60).  
 GN GLTSCR2.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OC NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.  
 RX MEDLINE=20175430; PubMed=10708517;  
 RA Smith J.S., Tachibana I., Pohl U., Lee H.K., Thanarajasingam U.,  
 Portier B.P., Ueki K., Billings S., Ramaswamy S., Mohrenweiser H.W.,  
 Scheithauer B.W., Louis D.N., Jenkins R.B.;  
 RT "A transcript map of the chromosome 19q-Arm glioma tumor suppressor  
 region.";  
 RL Genomics 64:44-50(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX SEQUENCE FROM N.A.  
 RC TISSUE=Muscle;  
 RA Strausberg R.;  
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE OF 9-478 FROM N.A., AND SUBCELLULAR LOCATION.  
 RX MEDLINE=99214318; PubMed=10196275;  
 RA Bruni R., Fineschi B., Ogle W.O., Roizman B.;  
 RT "A novel cellular protein, p60, interacting with both herpes simplex  
 virus 1 regulatory proteins ICP22 and ICP0 is modified in a  
 cell-type-specific manner and is recruited to the nucleus after  
 infection.";  
 RL J. Virol. 73:3810-3817(1999).  
 RN [4]  
 RP SEQUENCE OF 12-478 FROM N.A.  
 RA Andreu N., Estivill X., Escarceller M., Sunoy L.;  
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
 RN [5]  
 RP SEQUENCE OF 218-477 FROM N.A.  
 RC TISSUE=Testis;  
 RA Poustka A., Klein M., Mewes H.-W., Gassenhuber J., Wiemann S.;  
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.  
 CC -1- SUBUNIT: Interacts with HSV-1 early proteins ICP22 and ICP0.  
 CC -1- SUBCELLULAR LOCATION: Nuclear.  
 CC -1- TISSUE SPECIFICITY: Expressed at high levels in heart and  
 pancreas, moderate levels in placenta, liver, skeletal muscle, and  
 kidney, and low levels in brain and lung.  
 CC -1- SIMILARITY: BELONGS TO THE GLTSCR2 FAMILY.

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).  
 CC  
 CC EMBL; AF182076; AAF62873.1; -;  
 CC EMBL; BC004229; AAH04229.1; -;  
 CC EMBL; BC006311; AAH06311.1; -;  
 CC  
 DR EMBL; AF182076; AAF62873.1; -;  
 DR EMBL; BC004229; AAH04229.1; -;  
 DR EMBL; BC006311; AAH06311.1; -;  
 DR

RESULT 11  
 CC37\_SCHPO STANDARD; PRT; 466 AA.  
 ID CC37\_SCHPO STANDARD; PRT; 466 AA.  
 AC O94740;  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Hsp90 co-chaperone Cdc37 (Hsp90 chaperone protein kinase-targeting  
 subunit) (Cell division control protein 37).  
 GN CDC37 OR SPAC9B6.10.  
 OS Schizosaccharomyces pombe (Fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OC Schizosaccharomycetes.  
 GN NCBI\_TaxID=4896;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX Westwood P.K., Preston N.C., Fantes P.A.;  
 RT "Schizosaccharomycetes pombe cdc37 gene.";  
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX STRAIN=972;  
 RX MEDLINE=21848401; PubMed=11859360;  
 RA Wood V., Williams R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,  
 Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,  
 Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,  
 Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,  
 Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,  
 Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,  
 James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,  
 Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,  
 Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,  
 Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,  
 Skelton J., Simmonds M., Squares R., Squares S., Warren T., Whitehead S.,  
 Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
 Woodward J., Voiclaert G., Aert R., Robben J., Grymonprez B.,  
 Welljens I., Vansteelt E., Rieger M., Schaefer M., Mueller-Auer S.,  
 Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,  
 Raczyn K., Langer I., Beck A., Leirach H., Reinhardt R., Pohl T.M.,  
 Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,  
 Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,  
 Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
 Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Rhode G.,  
 Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
 Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
 Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,  
 Shpakovski G.V., Ussery D., Barrett B.G., Nurse P.;  
 RT "The genome sequence of Schizosaccharomycetes pombe.";  
 RL Nature 415:871-880(2002).  
 CC -1- FUNCTION: With Hsp90 it forms a complex that binds to several  
 kinases, resulting in stabilization and promotion of their  
 activity (By similarity).  
 CC -1- SUBUNIT: Forms a complex with Hsp90. Interacts with a number of  
 kinases (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -1- SIMILARITY: BELONGS TO THE CDC37 FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).  
 CC  
 CC EMBL; AJ132377; CAB38758.1; -;  
 CC EMBL; AJ132376; CAB38757.1; -;  
 CC EMBL; AL049769; CAB32371.2; -;  
 KW Chapterone; Cell division; Cell cycle.  
 SQ SEQUENCE 466 AA; 52554 MW; 647238B34CABB3C5 CRC64;

DR EMBL; BC010095; AAH10095.1; -  
 DR EMBL; AF296124; AAG30413.1; -  
 DR EMBL; AL359335; CAB94786.1; -  
 DR EMBL; AL359336; CAB94787.1; -  
 DR EMBL; AL122063; CAB59242.1; -  
 DR SWISS-2DPAGE; O9NZM5; HUMAN.  
 DR Genew; HGNC:4333; GLUTSCR2.  
 DR MIM; 605691; -  
 KW Nuclear protein; Polymorphism.  
 FT VARIANT 389 389 R -> Q.  
 FT /FTId=VAR\_011486.  
 FT GGS -> HEG (IN REF. 2; AAH04229).  
 FT G -> R (IN REF. 3).  
 FT RRKEQLWEKLAGKGLPREVRRAQAARLLNPSATRAKPGPOD  
 FT TVERP -> SGRSSYGRSWSPSRASSPGGAQGPSVAQPCN  
 FT KGNPAPAGHRIAA (IN REF. 3).  
 FT SDNPLDRPLVGQDEFFLE -> LNNPDKFVVPVPGCLFPG  
 FT (IN REF. 3).  
 FT A -> S (IN REF. 2; AAH04229).  
 FT D -> H (IN REF. 3).  
 FT PEGNLRDRFKSFFORNNMIEPRERAKFKRKYKVLVEKRAF  
 FT REIQ -> VLVVSCRGAPCPVTPSLPVPVPRGYGRHHGCP  
 FT WAGVGPMPRG (IN REF. 5).  
 FT EGNILDRFKSFFORNNMIEPRERAKFKRKYKVLVEKRAF  
 FT EIQ -> RGHSEFGSRAFRGGI (IN REF. 3).  
 FT SEQUENCE 478 AA; 54417 MW; 7F18923E348CB52B CRC64;  
 Query Match 63.5%; Score 33; DB 1; Length 478;  
 Best Local Similarity 60.0%; Pred. No. 30;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
 QY 2 EVVPXGMSYS 11  
 ID BCN5\_CLOPE STANDARD; PRT; 890 AA.  
 AC P08696;  
 DT 01-JAN-1988 (Rel. 06, Created)  
 DT 01-JAN-1988 (Rel. 06, Last sequence update)  
 DT 01-OCT-1994 (Rel. 30, Last annotation update)  
 DE Bacteriocin BCN5.  
 GN BCN.  
 OS Clostridium perfringens.  
 OG Plasmid pIP404.  
 OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;  
 OC Clostridium.  
 OX NCBI\_TaxID=1502;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CPN50;  
 RX MEDLINE=88336297; PubMed=2901769;  
 RA Garnier T., Cole S.T.;  
 RT "Complete nucleotide sequence and genetic organization of the  
 RT bacteriocinogenic plasmid, pIP404, from Clostridium perfringens.";  
 RL Plasmid 19:134-150(1988).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CPN50;  
 RX MEDLINE=87057020; PubMed=2877971;  
 RA Garnier T., Cole S.T.;  
 RT "Characterization of a bacteriocinogenic plasmid from Clostridium  
 RT perfringens and molecular genetic analysis of the  
 RT bacteriocin-encoding gene.";  
 RL J. Bacteriol. 168:1189-1196(1986).  
 RN [3]  
 RP SEQUENCE OF 1-14 FROM N.A.  
 RC STRAIN=CPN50;  
 RX MEDLINE=89039249; PubMed=2460717;  
 RA Garnier T., Cole S.T.;  
 RT "Studies of UV-inducible promoters from Clostridium perfringens in

RT vivo and in vitro".  
 RL Mol. Microbiol. 2:607-614(1988).  
 CC -1- FUNCTION: MAY FUNCTION AS AN IONOPHORE.  
 CC -1- INDUCTION: BY UV IRRADIATION.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M14481; AAA9248.1; -  
 DR EMBL; M32882; AAA98249.1; -  
 DR PIR; A30481; A30481.  
 DR InterPro; IPR003646; SH3\_bac.  
 DR SMART; SM00287; SH3b; 3.  
 KW Antibiotic; Bacteriocin; Plasmid.  
 FT DOMAIN 815 869 HYDROPHOBIC.  
 FT SEQUENCE 890 AA; 96699 MW; F4E5E8971C31C6C6 CRC64;  
 Query Match 63.5%; Score 33; DB 1; Length 890;  
 Best Local Similarity 66.7%; Pred. No. 58;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 QY 2 EVVPXGMSYS 10  
 ID LTBS\_HUMAN STANDARD; PRT; 1394 AA.  
 AC P22064;  
 DT 01-AUG-1991 (Rel. 19, Created)  
 DT 01-AUG-1991 (Rel. 19, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Latent transforming growth factor beta binding protein 1 precursor  
 DE (Transforming growth factor beta-1 binding protein 1) (TGF-beta1-BP-  
 DE 1).  
 GN LTBP1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RC TISSUE=Fibroblast, and Platelet;  
 RX MEDLINE=90275601; PubMed=2350783;  
 RA Kanazaki T., Olofsson A., Moren A., Wernstedt C., Hellman U.,  
 RA Miyazono K., Claesson-Welsh L., Heldin C.-H.;  
 RT "TGF-beta 1 binding protein: a component of the large latent complex  
 RT of TGF-beta 1 with multiple repeat sequences.";  
 RL Cell 61:1051-1061(1990).  
 CC -1- SUBUNIT: THE LARGE LATENT COMPLEX OF TGF-BETAL FROM PLATELETS IS  
 CC COMPOSED OF THE TGF-BETAL MOLECULE NONCOVALENTLY ASSOCIATED WITH  
 CC A DISULFIDE-BONDED COMPLEX OF A DIMER OF THE N-TERMINAL PROPEPTIDE  
 CC OF THE TGF-BETAL PRECURSOR AND A THIRD COMPONENT DENOTED TGF-  
 CC BETAL-BP. TGF-BETAL-BP DOES NOT BIND DIRECTLY TO ACTIVE TGF-BETAL.  
 CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; a short form (shown here) and a  
 CC long form (AC Q14766); are produced by alternative splicing.  
 CC -1- PTM: CONTAINS HYDROXYLATED ASPARAGINE RESIDUES.  
 CC -1- PTM: The N-terminus is blocked.  
 CC -1- SIMILARITY: CONTAINS 16 EGF-LIKE DOMAINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).



CC EMBL; M34057; AA61160.1; -  
DR PIR; A35626; A35626.  
DR HSP; P00750; ITRG.  
DR GlycoSultedB; P22064; -  
DR Genew; HGNC:6714; LTPB1.  
DR MIM; 150390;  
DR InterPro; IPR000152; Asx\_hydroxyl.  
DR InterPro; IPR000561; EGF-like.  
DR InterPro; IPR001881; EGF\_Ca.  
DR InterPro; IPR002212; Fibril-assoc.  
DR Pfam; PF00683; TB; 4.  
DR SMART; SM00179; EGF\_CA; 13.  
DR SMART; SM00001; EGF\_like; 4.  
DR PROSITE; PS00010; ASX\_HYDROXYL; 13.  
DR PROSITE; PS00022; EGF\_1; 2.  
DR PROSITE; PS01186; EGF\_2; 11.  
DR PROSITE; PS01187; EGF\_CA; 15.  
DR Growth factor binding; Repeat; EGF-like domain; Hydroxylation; Signal;  
KW Glycoprotein; Alternative splicing.  
KW SIGNAL  
FT CHAIN 1 20 POTENTIAL.  
FT 21 1394 LATENT TRANSFORMING GROWTH FACTOR BETA  
FT BINDING PROTEIN IS.  
FT EGF-LIKE 1, CALCIUM-BINDING (POTENTIAL).  
FT REPEAT A, 2, CALCIUM-BINDING (POTENTIAL).  
FT EGF-LIKE 3, CALCIUM-BINDING (POTENTIAL).  
FT EGF-LIKE 4, CALCIUM-BINDING (POTENTIAL).  
FT EGF-LIKE 5, CALCIUM-BINDING (POTENTIAL).  
FT EGF-LIKE 6, CALCIUM-BINDING (POTENTIAL).  
FT EGF-LIKE 7, CALCIUM-BINDING (POTENTIAL).  
FT EGF-LIKE 8, CALCIUM-BINDING (POTENTIAL).  
FT EGF-LIKE 9, CALCIUM-BINDING (POTENTIAL).  
FT EGF-LIKE 10, CALCIUM-BINDING (POTENTIAL).  
FT EGF-LIKE 11, CALCIUM-BINDING (POTENTIAL).  
FT EGF-LIKE 12, CALCIUM-BINDING (POTENTIAL).  
FT REPEAT B.  
FT EGF-LIKE 13, CALCIUM-BINDING (POTENTIAL).  
FT REPEAT C.  
FT EGF-LIKE 14, CALCIUM-BINDING (POTENTIAL).  
FT EGF-LIKE 15.  
FT EGF-LIKE 16, CALCIUM-BINDING (POTENTIAL).  
FT CELL ATTACHMENT SITE (POTENTIAL).  
FT SITE  
FT 304 315 BY SIMILARITY.  
FT DISULFID 310 324 BY SIMILARITY.  
FT DISULFID 326 339 BY SIMILARITY.  
FT DISULFID 550 562 BY SIMILARITY.  
FT DISULFID 557 571 BY SIMILARITY.  
FT DISULFID 573 586 BY SIMILARITY.  
FT DISULFID 592 604 BY SIMILARITY.  
FT DISULFID 599 613 BY SIMILARITY.  
FT DISULFID 615 628 BY SIMILARITY.  
FT DISULFID 634 645 BY SIMILARITY.  
FT DISULFID 640 654 BY SIMILARITY.  
FT DISULFID 657 669 BY SIMILARITY.  
FT DISULFID 675 686 BY SIMILARITY.  
FT DISULFID 681 695 BY SIMILARITY.  
FT DISULFID 698 709 BY SIMILARITY.  
FT DISULFID 715 726 BY SIMILARITY.  
FT DISULFID 721 735 BY SIMILARITY.  
FT DISULFID 737 750 BY SIMILARITY.  
FT DISULFID 756 767 BY SIMILARITY.  
FT DISULFID 762 776 BY SIMILARITY.  
FT DISULFID 778 791 BY SIMILARITY.  
FT DISULFID 797 808 BY SIMILARITY.  
FT DISULFID 803 817 BY SIMILARITY.  
FT DISULFID 819 832 BY SIMILARITY.  
FT DISULFID 838 850 BY SIMILARITY.  
FT DISULFID 845 859 BY SIMILARITY.  
FT DISULFID 861 873 BY SIMILARITY.  
FT DISULFID 879 891 BY SIMILARITY.  
FT DISULFID 885 900 BY SIMILARITY.

CC ENZYME WHICH IS COMPOSED OF 2 ALPHA CHAINS, 1 BETA CHAIN, AND 1  
 CC BETA' CHAIN (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE RNA POLYMERASE BETA' CHAIN FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; AE004121; AAF93502.1; -  
 CC HSP; O9KWU6; IHQM.  
 CC TIGR; VC0329; -  
 CC InterPro; IPR000722; RNA\_pol\_A.  
 CC InterPro; IPR002879; RNA\_pol\_A2.  
 CC Pfam; PF00623; RNA\_pol\_A; 1.  
 CC Pfam; PF01854; RNA\_pol\_A2; 2.  
 CC Transferase; DNA-directed RNA polymerase; Transcription;  
 CC Complete proteome.  
 CC SEQUENCE 1401 AA; 155021 MW; DFD80F2B5514504F CRC64;  
 CC  
 CC Query Match 63.5%; Score 33; DB 1; Length 1401;  
 CC Best Local Similarity 50.0%; Pred. No. 93;  
 CC Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
 CC  
 CC QY 2 EVVPXGMSYS 11  
 CC :|||:|  
 CC Db 581 QIVPKGLPVS 590  
 CC  
 CC RESULT 16  
 CC LTBL\_HUMAN  
 CC ID LTBL\_HUMAN STANDARD; PRT; 1595 AA.  
 CC AC Q14766;  
 CC DT 15-JUN-2002 (Rel. 41, Created)  
 CC DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 CC DE Latent transforming growth factor beta binding protein 1 (TGF-beta1-BP-  
 CC (Transforming growth factor beta-1 binding protein 1) (TGF-beta1-BP-  
 CC 1)).  
 CC GN LTBL.  
 CC OS Homo sapiens (Human).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 CC NCBI\_TaxID=9606;  
 CC [1]  
 CC RP SEQUENCE OF 1-346 FROM N.A.  
 CC RC TISSUE=Blood;  
 CC RX MEDLINE=96125117; PubMed=8537398;  
 CC RA Olofsson A., Ichijo H., Moren A., ten Dijke P., Miyazono K.,  
 CC Heldin C.-H.;  
 CC "Efficient association of an amino-terminally extended form of human  
 CC latent transforming growth factor-beta binding protein with the  
 CC extracellular matrix.";  
 CC J. Biol. Chem. 270:31294-31297(1995).  
 CC [2]  
 CC RP SEQUENCE OF 347-1595 FROM N.A.  
 CC RC TISSUE=Fibroblast, and Platelet;  
 CC RX MEDLINE=90275601; PubMed=2350783;  
 CC RA Kanazaki T., Olofsson A., Moren A., Wernstedt C., Hellman U.,  
 CC Miyazono K., Claesson-Welsh L., Heldin C.-H.;  
 CC "TGF-beta 1 binding protein: a component of the large latent complex  
 CC of TGF-beta 1 with multiple repeat sequences.";  
 CC Cell 61:1051-1061(1990).  
 CC -1- SUBUNIT: THE LARGE LATENT COMPLEX OF TGF-BETA1 FROM PLATELETS IS  
 CC A DISULFIDE-BONDED COMPLEX OF A DIMER OF THE N-TERMINAL PROPEPTIDE  
 CC OF THE TGF-BETA1 PRECURSOR AND A THIRD COMPONENT DENOTED TGF-  
 CC BETA1-BP. TGF-BETA1-BP DOES NOT BIND DIRECTLY TO ACTIVE TGF-BETA1.  
 CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; a short form (AC P22064) and a  
 CC long form (shown here); are produced by alternative splicing.

CC -1- TISSUE SPECIFICITY: The long isoform is found in fibroblasts.  
 CC -1- PTM: CONTAINS HYDROXYLATED ASPARAGINE RESIDUES (BY SIMILARITY).  
 CC -1- SIMILARITY: CONTAINS 16 EGF-LIKE DOMAINS.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; L48925; AAA96327.1; -  
 CC HSP; M34057; AAA61160.1; ALT\_INIT.  
 CC HSP; P08709; 1BF9.  
 CC MIM; 150390; -  
 CC InterPro; IPR000152; Asx\_hydroxyl.  
 CC InterPro; IPR000561; EGF-like.  
 CC InterPro; IPR001881; EGF-Ca.  
 CC InterPro; IPR002212; Fibril-assoc.  
 CC Pfam; PF00008; EGF; 15.  
 CC Pfam; PF00683; TB; 4.  
 CC SMART; SM00181; EGF; 1.  
 CC PROSITE; PS00010; ASX\_HYDROXYL; 13.  
 CC PROSITE; PS00022; EGF\_1; 2.  
 CC PROSITE; PS01186; EGF\_2; 11.  
 CC PROSITE; PS01187; EGF\_Ca; 15.  
 CC Growth factor binding; Repeat; EGF-like domain; Hydroxylation; Signal;  
 CC Glycoprotein; Alternative splicing.  
 CC SIGNAL 1 23 POTENTIAL.  
 CC CHAIN 24 1595  
 CC  
 CC DOMAIN 501 541  
 CC REPEAT EGF-LIKE 1, CALCIUM-BINDING (POTENTIAL).  
 CC REPEAT A. 549 613  
 CC DOMAIN 747 788  
 CC REPEAT EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).  
 CC DOMAIN 789 830  
 CC REPEAT EGF-LIKE 3, CALCIUM-BINDING (POTENTIAL).  
 CC DOMAIN 831 871  
 CC REPEAT EGF-LIKE 4, CALCIUM-BINDING (POTENTIAL).  
 CC DOMAIN 872 911  
 CC REPEAT EGF-LIKE 5, CALCIUM-BINDING (POTENTIAL).  
 CC DOMAIN 912 952  
 CC REPEAT EGF-LIKE 6, CALCIUM-BINDING (POTENTIAL).  
 CC DOMAIN 953 993  
 CC REPEAT EGF-LIKE 7, CALCIUM-BINDING (POTENTIAL).  
 CC DOMAIN 994 1034  
 CC REPEAT EGF-LIKE 8, CALCIUM-BINDING (POTENTIAL).  
 CC DOMAIN 1035 1075  
 CC REPEAT EGF-LIKE 9, CALCIUM-BINDING (POTENTIAL).  
 CC DOMAIN 1076 1117  
 CC REPEAT EGF-LIKE 10, CALCIUM-BINDING (POTENTIAL).  
 CC DOMAIN 1118 1159  
 CC REPEAT EGF-LIKE 11, CALCIUM-BINDING (POTENTIAL).  
 CC DOMAIN 1160 1202  
 CC REPEAT EGF-LIKE 12, CALCIUM-BINDING (POTENTIAL).  
 CC DOMAIN 1218 1285  
 CC REPEAT B. 1285 1340  
 CC REPEAT EGF-LIKE 13, CALCIUM-BINDING (POTENTIAL).  
 CC DOMAIN 1341 1463  
 CC REPEAT C. 1463 1535  
 CC DOMAIN 1535 1580  
 CC REPEAT EGF-LIKE 14, CALCIUM-BINDING (POTENTIAL).  
 CC DOMAIN 1580 1650  
 CC REPEAT EGF-LIKE 15, CALCIUM-BINDING (POTENTIAL).  
 CC SITE CELL ATTACHMENT SITE (POTENTIAL).  
 CC DISULFID 505 516  
 CC BY SIMILARITY.  
 CC DISULFID 511 525  
 CC BY SIMILARITY.  
 CC DISULFID 527 540  
 CC BY SIMILARITY.  
 CC DISULFID 751 763  
 CC BY SIMILARITY.  
 CC DISULFID 758 772  
 CC BY SIMILARITY.  
 CC DISULFID 774 787  
 CC BY SIMILARITY.  
 CC DISULFID 793 805  
 CC BY SIMILARITY.  
 CC DISULFID 800 814  
 CC BY SIMILARITY.  
 CC DISULFID 816 829  
 CC BY SIMILARITY.  
 CC DISULFID 835 846  
 CC BY SIMILARITY.  
 CC DISULFID 841 855  
 CC BY SIMILARITY.  
 CC DISULFID 858 870  
 CC BY SIMILARITY.  
 CC DISULFID 876 887  
 CC BY SIMILARITY.  
 CC DISULFID 882 896  
 CC BY SIMILARITY.  
 CC DISULFID 899 910  
 CC BY SIMILARITY.  
 CC DISULFID 916 927  
 CC BY SIMILARITY.  
 CC DISULFID 922 936  
 CC BY SIMILARITY.  
 CC DISULFID 938 951  
 CC BY SIMILARITY.  
 CC DISULFID 957 968  
 CC BY SIMILARITY.  
 CC DISULFID 963 977  
 CC BY SIMILARITY.  
 CC DISULFID 979 992  
 CC BY SIMILARITY.





DE	GN	CCND2 OR CYL-2.	DE	GN	CCND2 OR CYL-2.
OS	Mus musculus (Mouse).		OS	Mus musculus (Mouse).	
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;		OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.		OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.	
OX	NCBI_TaxID=10090;		OX	NCBI_TaxID=10090;	
RN	[1]		RN	[1]	
RP	SEQUENCE FROM N.A.		RP	SEQUENCE FROM N.A.	
RX	MEDLINE=92196134; PubMed=1372445;		RX	MEDLINE=92196134; PubMed=1372445;	
RA	Kiyokawa H., Busquets X., Powell C.T., Ngo L., Rifkind R.A.,		RA	Kiyokawa H., Busquets X., Powell C.T., Ngo L., Rifkind R.A.,	
RA	Marks P.A.;		RA	Marks P.A.;	
RT	"Cloning of a D-type cyclin from murine erythroleukemia cells.;"		RT	"Cloning of a D-type cyclin from murine erythroleukemia cells.;"	
RL	Proc. Natl. Acad. Sci. U.S.A. 89:2444-2447(1992).		RL	Proc. Natl. Acad. Sci. U.S.A. 89:2444-2447(1992).	
RN	[2]		RN	[2]	
RP	SEQUENCE FROM N.A.		RP	SEQUENCE FROM N.A.	
RX	MEDLINE=91235305; PubMed=1827757;		RX	MEDLINE=91235305; PubMed=1827757;	
RA	Matsushime H., Roussel M.F., Ashmun R.A., Sherr C.J.;		RA	Matsushime H., Roussel M.F., Ashmun R.A., Sherr C.J.;	
RT	"Colony-stimulating factor 1 regulates novel cyclins during the G1		RT	"Colony-stimulating factor 1 regulates novel cyclins during the G1	
RT	phase of the cell cycle.;"		RT	phase of the cell cycle.;"	
RL	Cell 65:701-713(1991).		RL	Cell 65:701-713(1991).	
CC	-1- FUNCTION: ESSENTIAL FOR THE CONTROL OF THE CELL CYCLE AT THE G1/S		CC	-1- FUNCTION: ESSENTIAL FOR THE CONTROL OF THE CELL CYCLE AT THE G1/S	
CC	(START) TRANSITION		CC	(START) TRANSITION	
CC	-1- SUBUNIT: INTERACTS WITH THE CDK4 AND CDK6 PROTEIN KINASES TO FORM		CC	-1- SUBUNIT: INTERACTS WITH THE CDK4 AND CDK6 PROTEIN KINASES TO FORM	
CC	A SERINE/THREONINE KINASE HOLOENZYME COMPLEX. THE CYCLIN SUBUNIT		CC	A SERINE/THREONINE KINASE HOLOENZYME COMPLEX. THE CYCLIN SUBUNIT	
CC	IMPARTS SUBSTRATE SPECIFICITY TO THE COMPLEX.		CC	IMPARTS SUBSTRATE SPECIFICITY TO THE COMPLEX.	
CC	-1- SIMILARITY: BELONGS TO THE CYCLIN FAMILY. CYCLIN D SUBFAMILY.		CC	-1- SIMILARITY: BELONGS TO THE CYCLIN FAMILY. CYCLIN D SUBFAMILY.	
CC	-----		CC	-----	
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		CC	This SWISS-PROT entry is copyright. It is produced through a collaboration	
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -	
CC	the European Bioinformatics Institute. There are no restrictions on its		CC	the European Bioinformatics Institute. There are no restrictions on its	
CC	use by non-profit institutions as long as its content is in no way		CC	use by non-profit institutions as long as its content is in no way	
CC	modified and this statement is not removed. Usage by and for commercial		CC	modified and this statement is not removed. Usage by and for commercial	
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>		CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>	
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).		CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).	
CC	-----		CC	-----	
DR	EMBL; M83749; AAA37519.1; -		DR	EMBL; M83749; AAA37519.1; -	
DR	EMBL; M86182; AAA37503.1; -		DR	EMBL; M86182; AAA37503.1; -	
DR	PIR; B40035; B40035.		DR	PIR; B40035; B40035.	
DR	PIR; A41984; A41984.		DR	PIR; A41984; A41984.	
DR	MGI; 88314; Ccnd2.		DR	MGI; 88314; Ccnd2.	
DR	InterPro; IPR004366; Cyclin.		DR	InterPro; IPR004366; Cyclin.	
DR	InterPro; IPR004367; Cyclin_Cterm.		DR	InterPro; IPR004367; Cyclin_Cterm.	
DR	Pfam; PF00134; cyclin; 1.		DR	Pfam; PF00134; cyclin; 1.	
DR	Pfam; PF02984; cyclin_C; 1.		DR	Pfam; PF02984; cyclin_C; 1.	
DR	SMART; SM00385; CYCLIN; 1.		DR	SMART; SM00385; CYCLIN; 1.	
DR	PROSITE; PS00292; CYCLINS; 1.		DR	PROSITE; PS00292; CYCLINS; 1.	
DR	Cyclin; Cell cycle; Cell division; Multigene family.		DR	Cyclin; Cell cycle; Cell division; Multigene family.	
KW	Cyclin; Cell cycle; Cell division; Multigene family.		KW	Cyclin; Cell cycle; Cell division; Multigene family.	
SQ	SEQUENCE 289 AA; 32897 MW; 58F322771DD1DA3D CRC64;		SQ	SEQUENCE 289 AA; 32897 MW; 58F322771DD1DA3D CRC64;	
Query Match	61.5%; Score 32; DB 1; Length 289;		Query Match	61.5%; Score 32; DB 1; Length 289;	
Best Local Similarity	60.0%; Pred. No. 29;		Best Local Similarity	60.0%; Pred. No. 29;	
Matches	6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;		Matches	6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;	
QY	1 EEVVPXGMSY 10		QY	1 EEVVPXGMSY 10	
Db			Db		
	73 EEVFPPLAMNY 82			73 EEVFPPLAMNY 82	
RESULT 21			RESULT 21		
CGD1_BRARE	STANDARD; PRT; 291 AA.		CGD1_BRARE	STANDARD; PRT; 291 AA.	
ID	CGD1_BRARE		ID	CGD1_BRARE	
AC	Q90459;		AC	Q90459;	
DT	15-JUL-1998 (Rel. 36, Created)		DT	15-JUL-1998 (Rel. 36, Created)	
DT	15-JUL-1998 (Rel. 36, Last sequence update)		DT	15-JUL-1998 (Rel. 36, Last sequence update)	
DT	16-OCT-2001 (Rel. 40, Last annotation update)		DT	16-OCT-2001 (Rel. 40, Last annotation update)	
DE	GL/S-specific cyclin D1.		DE	GL/S-specific cyclin D1.	
GN	CYCDL.		GN	CYCDL.	
OS	Brachydanio rerio (Zebrafish) (Danio rerio).		OS	Brachydanio rerio (Zebrafish) (Danio rerio).	
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;		OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;	
OC	Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;		OC	Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;	
OC	Cyprinidae; Danio.		OC	Cyprinidae; Danio.	
OX	NCBI_TaxID=7955;		OX	NCBI_TaxID=7955;	
RN	[1]		RN	[1]	

```

RP SEQUENCE FROM N.A.
RX MEDLINE-96138542; PubMed-8547308;
RA Yarden A., Salomon D., Geiger B.;
RT "zebrafish cyclin D1 is differentially expressed during early
RT embryogenesis.";
RL Biochim. Biophys. Acta 1264:257-260(1995).
CC -1- FUNCTION: ESSENTIAL FOR THE CONTROL OF THE CELL CYCLE AT THE G1/S
CC (START) TRANSITION.
CC -1- SUBUNIT: INTERACTS WITH THE CDK4 AND CDK6 PROTEIN KINASES TO FORM
CC A SERINE/THREONINE KINASE HOLOENZYME COMPLEX. THE CYCLIN SUBUNIT
CC IMPARTS SUBSTRATE SPECIFICITY TO THE COMPLEX.
CC -1- SIMILARITY: BELONGS TO THE CYCLIN FAMILY. CYCLIN D SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: X87581; CAA60885.1;
CC ZFIN: ZDB-GENE-980526-176; cyclin.
CC InterPro: IPR004366; Cyclin.
CC Pfam: PF00134; cyclin; 1.
CC SMART: PF02984; cyclin_C; 1.
CC SMART: SM00385; CYCLIN; 1.
CC PROSITE: PS00292; CYCLINS; 1.
CC CYCLIN; Cell cycle; Cell division.
CC SEQUENCE 291 AA; 33067 MW; FA5274CB1B46D5EF CRC64;
CC
CC Query Match 61.5%; Score 32; DB 1; Length 291;
CC Best Local Similarity 60.0%; Pred. No. 29;
CC Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
CC
CC QY 1 EEVVPXGMSY 10
CC ||| | | |
CC Db 75 EEVFPPLAMNY 84
CC
CC RESULT 22
CC GCD1_XENLA STANDARD; PRT; 291 AA.
CC AC P30755;
CC DT 01-OCT-1996 (Rel. 34, Created)
CC DT 01-OCT-1996 (Rel. 34, Last sequence update)
CC DT 16-OCT-2001 (Rel. 40, Last annotation update)
CC DE G1/S-specific cyclin D1.
CC GN CCND1.
CC OS Xenopus laevis (African clawed frog).
CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
CC OC Xenopodinae; Xenopus.
CC OX NCBI_TaxID=8355;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RA Cockerill M.J., Hunt T.;
CC RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: ESSENTIAL FOR THE CONTROL OF THE CELL CYCLE AT THE G1/S
CC (START) TRANSITION.
CC -1- SUBUNIT: INTERACTS WITH THE CDK4 AND CDK6 PROTEIN KINASES TO FORM
CC A SERINE/THREONINE KINASE HOLOENZYME COMPLEX. THE CYCLIN SUBUNIT
CC IMPARTS SUBSTRATE SPECIFICITY TO THE COMPLEX.
CC -1- SIMILARITY: BELONGS TO THE CYCLIN FAMILY. CYCLIN D SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC
CC QY 1 EEVVPXGMSY 10
CC ||| | | |
CC Db 75 EEVFPPLAMNY 84
CC
CC RESULT 23
CC CGD2_CHICK STANDARD; PRT; 291 AA.
CC AC P49706;
CC DT 01-OCT-1996 (Rel. 34, Created)
CC DT 01-OCT-1996 (Rel. 34, Last sequence update)
CC DT 16-OCT-2001 (Rel. 40, Last annotation update)
CC DE G1/S-specific cyclin D2.
CC GN CCND2.
CC OS Gallus gallus (Chicken).
CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC OC Archosauia; Aves; Neognathae; Galliformes; Phasianidae;
CC OC Gallus.
CC OX NCBI_TaxID=9031;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RX MEDLINE-96144302; PubMed-8566807;
CC RA Li H., Grenet J., Kidd V.J.;
CC RT "Structure and gene expression of avian cyclin D2.";
CC RL Gene 167:341-342(1995).
CC -1- FUNCTION: ESSENTIAL FOR THE CONTROL OF THE CELL CYCLE AT THE G1/S
CC (START) TRANSITION.
CC -1- SUBUNIT: INTERACTS WITH THE CDK4 AND CDK6 PROTEIN KINASES TO FORM
CC A SERINE/THREONINE KINASE HOLOENZYME COMPLEX. THE CYCLIN SUBUNIT
CC IMPARTS SUBSTRATE SPECIFICITY TO THE COMPLEX.
CC -1- SIMILARITY: BELONGS TO THE CYCLIN FAMILY. CYCLIN D SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: U28980; AAA96955.1;
CC InterPro: IPR004366; Cyclin.
CC InterPro: IPR004367; Cyclin_Cterm.
CC Pfam: PF00134; cyclin; 1.
CC Pfam: PF02984; cyclin_C; 1.
CC SMART: SM00385; CYCLIN; 1.
CC PROSITE: PS00292; CYCLINS; 1.
CC CYCLIN; Cell cycle; Cell division; Multigene family.
CC SEQUENCE 291 AA; 33163 MW; 43A7E646AEAF3109 CRC64;
CC
CC Query Match 61.5%; Score 32; DB 1; Length 291;
CC Best Local Similarity 60.0%; Pred. No. 29;
CC Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
CC
CC QY 1 EEVVPXGMSY 10
CC ||| | | |
CC Db 74 EEVFPPLAMNY 83

```

Wed Jun 11 15:42:04 2003

RESULT 24  
CGD2\_XENLA STANDARD; PRT; 291 AA.  
AC P53782;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE G1/S-specific cyclin D2.  
GN CCND2  
OS Xenopus laevis (African clawed frog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;  
OC Xenopodinae; Xenopus.  
OX NCBI\_TaxID=83355;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Cockerill M.J., Hunt T.;  
RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97380591; PubMed=9237366;  
RA Taleb F., Jessus C.;  
RT "Xenopus cyclin D2: cloning and expression in oocytes and during  
early development.";  
RT Biol. Cell 88:99-111(1996).  
RL BIOL. CELL 88:99-111(1996).  
CC -1- FUNCTION: ESSENTIAL FOR THE CONTROL OF THE CELL CYCLE AT THE G1/S  
(START) TRANSITION.  
CC -1- SUBUNIT: INTERACTS WITH THE CDK4 AND CDK6 PROTEIN KINASES TO FORM  
A SERINE/THREONINE KINASE HOLOENZYME COMPLEX. THE CYCLIN SUBUNIT  
IMPARTS SUBSTRATE SPECIFICITY TO THE COMPLEX.  
CC -1- SIMILARITY: BELONGS TO THE CYCLIN FAMILY. CYCLIN D SUBFAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL; U40844; AAA83271.1; -  
DR InterPro: IPR004366; Cyclin.  
DR InterPro: IPR004367; Cyclin\_Cterm.  
DR Pfam; PF00134; cyclin; 1.  
DR Pfam; PF02984; cyclin\_C; 1.  
DR SMART; SM00385; CYCLIN; 1.  
DR PROSITE; PS00292; CYCLINS; 1.  
DR Cyclin; Cell cycle; Cell division; Multigene family.  
KW CYCLIN; Cell cycle; Cell division; Multigene family.  
SQ SEQUENCE 292 AA; 33263 MW; 7B543029DB45A67D CRC64;  
-----  
Query Match 61.5%; Score 32; DB 1; Length 291;  
Best Local Similarity 60.0%; Pred. No. 30;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
QY 1 EEVVPXGMSY 10  
| | | | | | | |  
Db 75 EEVFPAMNY 84  
-----  
Search completed: June 10, 2003, 13:40:20  
Job time : 4.5 secs

RESULT 25  
CGDI\_CHICK STANDARD; PRT; 292 AA.  
AC P55169;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE G1/S-specific cyclin D1.  
GN CCND1.  
OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
-----  
Query Match 61.5%; Score 32; DB 1; Length 291;  
Best Local Similarity 60.0%; Pred. No. 29;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
QY 1 EEVVPXGMSY 10  
| | | | | | | |  
Db 74 EEVFPAMNY 83  
-----  
CGDI\_CHICK STANDARD; PRT; 292 AA.  
AC P55169;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE G1/S-specific cyclin D1.  
GN CCND1.  
OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2003, 13:30:25 ; Search time 25.7857 Seconds  
(without alignments)  
87.898 Million cell updates/sec

Title: US-09-909-164-10

Perfect score: 52

Sequence: 1 EEVVPXGMSYS 11

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 75 summaries

Database :

SPTREMBL.21.\*

1: sp.archaea.\*

2: sp.bacteria.\*

3: sp.fungi.\*

4: sp.human.\*

5: sp.invertebrate.\*

6: sp.mammal.\*

7: sp.mhc.\*

8: sp.organelle.\*

9: sp.phage.\*

10: sp.plant.\*

11: sp.rodent.\*

12: sp.virus.\*

13: sp.vertibrate.\*

14: sp.unclassified.\*

15: sp.virus.\*

16: sp.bacteriaph.\*

17: sp.archaeap.\*

pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	38	73.1	387	16	Q98FX1 rhizobium 1
2	38	73.1	1063	16	Q8RG86 fusobacteri
3	38	73.1	3472	1	O74056 cenarchaeum
4	37	71.2	840	3	Q9URY8 schizosacch
5	36	69.2	471	11	Q8RI26 mus musculu
6	36	69.2	484	11	Q8VD18 mus musculu
7	35	67.3	225	10	Q40129 lycopersico
8	35	67.3	425	5	Q9XVK4 caenorhabdi
9	35	67.3	556	4	O43733 homo sapien
10	35	67.3	583	5	Q9BHA5 plasmodium
11	35	67.3	583	5	Q9BH83 plasmodium
12	35	67.3	670	11	Q01487 rattus norv
13	35	67.3	749	16	Q9PDM6 xylella fas
14	35	67.3	1902	4	Q14122 homo sapien
15	34	65.4	156	3	Q12479 saccharomyc
16	34	65.4	219	17	Q971S2 sulfolobus

17	34	65.4	252	17	O28342	O28342 archaeoglob
18	34	65.4	290	4	Q96MU1	Q96MU1 homo sapien
19	34	65.4	387	16	Q92MD6	Q92MD6 rhizobium m
20	34	65.4	541	16	Q98BP5	Q98BP5 rhizobium 1
21	34	65.4	544	16	Q9PQD2	Q9PQD2 ureaplasma
22	34	65.4	842	3	Q9URR4	Q9URR4 penicillium
23	34	65.4	1049	16	Q8XT05	Q8XT05 ralstonia s
24	34	65.4	1499	4	Q969I4	Q969I4 homo sapien
25	33	63.5	143	17	Q8TX62	Q8TX62 methanopyru
26	33	63.5	162	11	Q9CXQ4	Q9CXQ4 mus musculu
27	33	63.5	165	17	O28330	O28330 archaeoglob
28	33	63.5	193	2	Q8VUA8	Q8VUA8 lactococcus
29	33	63.5	209	16	Q8RE56	Q8RE56 fusobacteri
30	33	63.5	284	16	P74187	P74187 synechocyst
31	33	63.5	298	10	Q9M3C0	Q9M3C0 arabidopsis
32	33	63.5	326	12	Q9Q9Q9	Q9Q9Q9 soil-borne
33	33	63.5	326	12	Q9Q9Q5	Q9Q9Q5 soil-borne
34	33	63.5	326	12	Q9QCE7	Q9QCE7 soil-borne
35	33	63.5	326	12	Q9DJG4	Q9DJG4 soil-borne
36	33	63.5	326	12	Q91DN1	Q91DN1 soil-borne
37	33	63.5	326	12	Q9Q9Q7	Q9Q9Q7 soil-borne
38	33	63.5	327	12	Q96360	Q96360 soil-borne
39	33	63.5	332	10	Q9FNL4	Q9FNL4 arabidopsis
40	33	63.5	368	16	Q9XOU3	Q9XOU3 thermotoga
41	33	63.5	393	5	Q9V9I4	Q9V9I4 drosophila
42	33	63.5	479	4	Q96CS0	Q96CS0 homo sapien
43	33	63.5	548	11	Q9D2X9	Q9D2X9 mus musculu
44	33	63.5	584	16	Q8R8K6	Q8R8K6 thermoaer
45	33	63.5	648	4	Q96MB2	Q96MB2 homo sapien
46	33	63.5	653	16	Q9KVE3	Q9KVE3 vibrio chol
47	33	63.5	676	5	Q9VA55	Q9VA55 drosophila
48	33	63.5	676	5	Q8T8Z7	Q8T8Z7 drosophila
49	33	63.5	678	12	Q9ELX6	Q9ELX6 cercopithec
50	33	63.5	746	3	Q9URR3	Q9URR3 penicillium
51	33	63.5	791	4	Q9H2K5	Q9H2K5 homo sapien
52	33	63.5	793	4	Q9H2K6	Q9H2K6 homo sapien
53	33	63.5	844	11	Q922D4	Q922D4 mus musculu
54	33	63.5	1028	16	Q8YJ11	Q8YJ11 brucella me
55	33	63.5	1152	16	Q9CC95	Q9CC95 mycobacteri
56	33	63.5	1305	5	Q9V7C7	Q9V7C7 drosophila
57	33	63.5	1394	4	Q8TD95	Q8TD95 homo sapien
58	33	63.5	1442	17	Q96YH5	Q96YH5 sulfolobus
59	33	63.5	1548	10	O65531	O65531 arabidopsis
60	33	63.5	1713	11	O88349	O88349 mus musculu
61	32	61.5	84	16	Q97DE7	Q97DE7 clostridium
62	32	61.5	103	11	Q9PDH9	Q9PDH9 mus musculu
63	32	61.5	108	1	Q9UX33	Q9UX33 sulfolobus
64	32	61.5	153	13	P79919	P79919 xenopus lae
65	32	61.5	156	11	Q9D8L9	Q9D8L9 mus musculu
66	32	61.5	174	10	Q9M3T4	Q9M3T4 betula verr
67	32	61.5	190	13	O57481	O57481 stizostedio
68	32	61.5	191	11	O99NB4	O99NB4 rattus norv
69	32	61.5	200	17	Q97CD0	Q97CD0 thermoplas
70	32	61.5	207	2	Q47284	Q47284 escherichia
71	32	61.5	234	2	O32330	O32330 clostridium
72	32	61.5	236	10	Q9SXF1	Q9SXF1 arabidopsis
73	32	61.5	240	11	Q9DB09	Q9DB09 mus musculu
74	32	61.5	243	12	Q91EW1	Q91EW1 cydia pomon
75	32	61.5	279	17	Q9Y8Z4	Q9Y8Z4 aeropyrum p

## ALIGNMENTS

RESULT 1  
Q98FX1 PRELIMINARY; PRT; 387 AA.  
ID Q98FX1  
AC Q98FX1  
DT 01-OCT-2001 (TrEMBLrel. 18, Created)  
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)  
DE Hippurate hydrolase.  
GN MLR3583.

OS Rhizobium loti (Mesorhizobium loti).  
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
 OC Phyllobacteriaceae; Mesorhizobium.  
 OX NCBI\_TaxID=381;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MAFF303099;  
 RX MEDLINE=21082930; PubMed=11214968;  
 RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,  
 RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,  
 RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,  
 RA Mochizuki Y., Nakayama S., Nakazaki N., Shimo S., Sugimoto M.,  
 RA Takeuchi C., Yamada M., Tabata S.;  
 RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium  
 RT Mesorhizobium loti.";  
 RL DNA Res. 7:331-338 (2000).  
 DR EMBL; AF003002; BAB50445.1; -;  
 DR InterPro; IPR002933; Peptidase\_M20.  
 DR Pfam; PF01546; Peptidase\_M20; 1.  
 KW Hydrolase; Complete proteome.  
 SQ SEQUENCE 387 AA; 41180 MW; 131BFF8E64306829 CRC64;

Query Match 73.1%; Score 38; DB 16; Length 387;  
 Best Local Similarity 60.0%; Pred. No. 9.8;  
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 EVVVPXGMSY 10  
 : : : : :  
 Db 367 DEAIHGMSY 376

## RESULT 2

OS Rhizobium loti (Mesorhizobium loti).  
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
 OC Phyllobacteriaceae; Mesorhizobium.  
 OX NCBI\_TaxID=381;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MAFF303099;  
 RX MEDLINE=21082930; PubMed=11214968;  
 RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,  
 RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,  
 RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,  
 RA Mochizuki Y., Nakayama S., Nakazaki N., Shimo S., Sugimoto M.,  
 RA Takeuchi C., Yamada M., Tabata S.;  
 RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium  
 RT Mesorhizobium loti.";  
 RL DNA Res. 7:331-338 (2000).  
 DR EMBL; AF003002; BAB50445.1; -;  
 DR InterPro; IPR002933; Peptidase\_M20.  
 DR Pfam; PF01546; Peptidase\_M20; 1.  
 KW Hydrolase; Complete proteome.  
 SQ SEQUENCE 387 AA; 41180 MW; 131BFF8E64306829 CRC64;

Query Match 73.1%; Score 38; DB 16; Length 1063;  
 Best Local Similarity 60.0%; Pred. No. 29;  
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 EVVVPXGMSY 11  
 : : : : :  
 Db 195 EIVPGLNYS 204

## RESULT 3

OS Rhizobium loti (Mesorhizobium loti).  
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
 OC Phyllobacteriaceae; Mesorhizobium.  
 OX NCBI\_TaxID=381;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MAFF303099;  
 RX MEDLINE=21082930; PubMed=11214968;  
 RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,  
 RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,  
 RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,  
 RA Mochizuki Y., Nakayama S., Nakazaki N., Shimo S., Sugimoto M.,  
 RA Takeuchi C., Yamada M., Tabata S.;  
 RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium  
 RT Mesorhizobium loti.";  
 RL DNA Res. 7:331-338 (2000).  
 DR EMBL; AF003002; BAB50445.1; -;  
 DR InterPro; IPR002933; Peptidase\_M20.  
 DR Pfam; PF01546; Peptidase\_M20; 1.  
 KW Hydrolase; Complete proteome.  
 SQ SEQUENCE 387 AA; 41180 MW; 131BFF8E64306829 CRC64;

Query Match 73.1%; Score 38; DB 16; Length 1063;  
 Best Local Similarity 60.0%; Pred. No. 29;  
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 EVVVPXGMSY 11  
 : : : : :  
 Db 195 EIVPGLNYS 204

DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE Hypothetical 367.1 kDa protein.  
 OS Cenarchaeum symbiosum.  
 OC Archaea; Crenarchaeota; Cenarchaeum.  
 OX NCBI\_TaxID=46770;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=B;  
 RX MEDLINE=98422450; PubMed=9748430;  
 RA Schleper C., DeLong E.F., Preston C.M., Feldman R.A., Wu K.Y.,  
 RA Swanson R.V.;  
 RT "Genomic analysis reveals chromosomal variation in natural populations  
 RT of the uncultured psychrophilic archaeon Cenarchaeum symbiosum.";  
 RL J. Bacteriol. 180:5003-5009 (1998).  
 CC 1- SIMILARITY: CONTAINS 5 WD REPEATS (TRP-ASP DOMAINS).  
 DR EMBL; AF083072; AAC62699.1; -;  
 DR InterPro; IPR000515; BPD\_transp.  
 DR Pfam; PF001680; WD40.  
 DR Pfam; PF00400; WD40; 4.  
 DR SMART; SM00320; WD40; 2.  
 DR PROSITE; PS00402; BPD\_TRANS\_INN\_MEMBR; UNKNOWN\_1.  
 KW Hypothetical protein; Repeat; WD repeat.  
 SQ SEQUENCE 3472 AA; 367058 MW; 37F80707030F9355 CRC64;

Query Match 73.1%; Score 38; DB 1; Length 3472;  
 Best Local Similarity 54.5%; Pred. No. 1e+02;  
 Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 BEVVPXGMSYS 11  
 : : : : :  
 Db 2294 EDVIPRGISFS 2304

## RESULT 4

OS Schizosaccharomyces pombe (Fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OC Schizosaccharomycetes.  
 OX NCBI\_TaxID=4896;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=972H-;  
 RA Hunt C., Aves S., McDougall R.C., Rajandream M.A., Barrell B.G.;  
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AL132779; CAB60015.1; -;  
 DR InterPro; IPR002645; STAS.  
 DR InterPro; IPR001902; Sulfate\_transp.  
 DR Pfam; PF01740; STAS; 1.  
 DR Pfam; PF00916; Sulfate\_transp; 1.  
 DR TIGRfams; TIGR00815; sulp; 1.  
 SQ SEQUENCE 840 AA; 93517 MW; ED4833E162B69077 CRC64;

Query Match 71.2%; Score 37; DB 3; Length 840;  
 Best Local Similarity 77.8%; Pred. No. 37;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMSYS 11  
 : : : : :  
 Db 135 VVPQMSYA 143

## RESULT 5

OS Schizosaccharomyces pombe (Fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OC Schizosaccharomycetes.  
 OX NCBI\_TaxID=4896;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=972H-;  
 RA Hunt C., Aves S., McDougall R.C., Rajandream M.A., Barrell B.G.;  
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AL132779; CAB60015.1; -;  
 DR InterPro; IPR002645; STAS.  
 DR InterPro; IPR001902; Sulfate\_transp.  
 DR Pfam; PF01740; STAS; 1.  
 DR Pfam; PF00916; Sulfate\_transp; 1.  
 DR TIGRfams; TIGR00815; sulp; 1.  
 SQ SEQUENCE 840 AA; 93517 MW; ED4833E162B69077 CRC64;

Query Match 71.2%; Score 37; DB 3; Length 840;  
 Best Local Similarity 77.8%; Pred. No. 37;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMSYS 11  
 : : : : :  
 Db 135 VVPQMSYA 143

## RESULT 5

OS Schizosaccharomyces pombe (Fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OC Schizosaccharomycetes.  
 OX NCBI\_TaxID=4896;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=972H-;  
 RA Hunt C., Aves S., McDougall R.C., Rajandream M.A., Barrell B.G.;  
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AL132779; CAB60015.1; -;  
 DR InterPro; IPR002645; STAS.  
 DR InterPro; IPR001902; Sulfate\_transp.  
 DR Pfam; PF01740; STAS; 1.  
 DR Pfam; PF00916; Sulfate\_transp; 1.  
 DR TIGRfams; TIGR00815; sulp; 1.  
 SQ SEQUENCE 840 AA; 93517 MW; ED4833E162B69077 CRC64;

Query Match 71.2%; Score 37; DB 3; Length 840;  
 Best Local Similarity 77.8%; Pred. No. 37;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMSYS 11  
 : : : : :  
 Db 135 VVPQMSYA 143

AC O8RL26;  
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE Hypothetical 54.5 kDa protein (Fragment).  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=LIVER;  
 RA Strausberg R.;  
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC025810; AAH25810.1; -;  
 KW Hypothetical protein.  
 FT NON\_TER 1  
 SQ SEQUENCE 471 AA; 54506 MW; E0DA685C374A9760 CRC64;  
 Query Match 69.2%; Score 36; DB 11; Length 471;  
 Best Local Similarity 60.0%; Pred. No. 32;  
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
 QY 2 EVVPGXGMSYS 11  
 Db 226 EVIPAGASYN 235  
 RESULT 6  
 Q8VD18 PRELIMINARY; PRT; 484 AA.  
 AC Q8VD18;  
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE Similar to glioma tumor suppressor candidate region gene 2.  
 GN AW536441.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=SALIVARY GLAND;  
 RA Strausberg R.;  
 RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC017637; AAH17637.1; -;  
 DR MGD; MGI:2138595; AW536441.  
 SQ SEQUENCE 484 AA; 55835 MW; BBB45F3B4BE02A36 CRC64;  
 Query Match 69.2%; Score 36; DB 11; Length 484;  
 Best Local Similarity 60.0%; Pred. No. 33;  
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
 QY 2 EVVPGXGMSYS 11  
 Db 239 EVIPAGASYN 248  
 RESULT 7  
 Q40129 PRELIMINARY; PRT; 225 AA.  
 AC Q40129;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 DE Hypothetical 25.2 kDa protein precursor.  
 OS Lycopersicon esculentum (Tomato).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.  
 OX NCBI\_TaxID=4081;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RC STRAIN=VF36; TISSUE=PISTIL;  
 RA MEDLINE-95375233; PubMed-7647301;  
 RA Milligan S.B.; Gasser C.S.;  
 RT "Nature and regulation of pistil-expressed genes in tomato.";  
 RL Plant Mol. Biol. 28:691-711(1995).  
 DR EMBL; U20592; AAA80497.1; -;  
 DR InterPro; IPR002160; Kunitz\_legume.  
 DR Pfam; PF00197; Kunitz\_legume; 1.  
 DR PRODOM; PD000891; Kunitz\_legume; 1.  
 DR SMART; SM00452; STI; 1.  
 DR PROSITE; PS00283; SOYBEAN\_KUNITZ; UNKNOWN\_1.  
 KW Hypothetical protein; Signal.  
 FT SIGNAL 1 20 POTENTIAL.  
 CHAIN 21 225 UNKNOWN.  
 SQ SEQUENCE 225 AA; 25188 MW; 1074C261D20CFDAD CRC64;  
 Query Match 67.3%; Score 35; DB 10; Length 225;  
 Best Local Similarity 54.5%; Pred. No. 24;  
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 EEVVPXGMSYS 11  
 Db 32 DEVPNGKTYA 42  
 RESULT 8  
 Q9XVK4 PRELIMINARY; PRT; 425 AA.  
 AC Q9XVK4;  
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
 DE 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 DE R10D12.10 protein.  
 GN R10D12.10.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Percy C.M.;  
 RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-99069613; PubMed-9851916;  
 RA none;  
 RT "Genome sequence of the nematode C.elegans: A platform for  
 RL investigating biology.";  
 RL Science 282:2012-2018(1998).  
 DR EMBL; Z81109; CAB03241.1; -;  
 DR InterPro; IPR000719; Euk\_pkinase.  
 DR Pfam; PF00069; pkinase; 1.  
 DR PRODOM; PD000001; Euk\_pkinase; 1.  
 DR PROSITE; PS0011; PROTEIN\_KINASE\_DOM; 1.  
 KW ATP-binding; Transferase.  
 SQ SEQUENCE 425 AA; 49410 MW; 5D96E29B08C8E9D6 CRC64;  
 Query Match 67.3%; Score 35; DB 5; Length 425;  
 Best Local Similarity 50.0%; Pred. No. 48;  
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 EEVVPXGMSY 10  
 Db 335 EQIVPGGLQY 344  
 RESULT 9  
 O43733 PRELIMINARY; PRT; 556 AA.  
 ID O43733  
 AC O43733;  
 DT 01-JUN-1998 (TrEMBLrel. 06, Created)  
 DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)

DT 01-JUN-1998 (TREMBlrel. 06, Last annotation update)  
 DE DNA binding protein (Fragment).  
 GN DJ451B15.2.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 ON NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Tubby B.;  
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; Z98050; CAB10847.1; -.  
 FT NON\_TER  
 SQ SEQUENCE 556 AA; 59059 MW; ECB00E403FB2528 CRC64;

Query Match 67.3%; Score 35; DB 4; Length 556;  
 Best Local Similarity 66.7%; Pred. No. 64;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 QY 3 VVPXGMSYS 11  
 Db 244 VVPAGLTYS 252

RESULT 10  
 Q9BHA5 PRELIMINARY; PRT; 583 AA.  
 AC Q9BHA5;  
 DT 01-JUN-2001 (TREMBlrel. 17, Created)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE Choline transporter.  
 GN SCT1.  
 OS Plasmodium falciparum.  
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
 ON NCBI\_TaxID=5833;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Ben Mamoun C., Gluzman I.Y., Goldberg D.E.;  
 RL "Plasmodium falciparum choline transporter (PfSCT1) gene."  
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY007372; AAK14816.1; -.  
 DR EMBL; AY007375; AAG17947.1; -.  
 DR InterPro: IPR002123; Acyltransferase.  
 DR Pfam: PF01553; Acyltransferase; 1.  
 SQ SEQUENCE 583 AA; 66917 MW; 2B2BFAE3E395E049 CRC64;

Query Match 67.3%; Score 35; DB 5; Length 583;  
 Best Local Similarity 55.6%; Pred. No. 67;  
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
 QY 3 VVPXGMSYS 11  
 Db 227 IIPVGLSYS 235

RESULT 11  
 Q9BH83 PRELIMINARY; PRT; 583 AA.  
 AC Q9BH83;  
 DT 01-JUN-2001 (TREMBlrel. 17, Created)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE Choline transporter.  
 GN SCT1.  
 OS Plasmodium falciparum.  
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
 ON NCBI\_TaxID=5833;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Ben Mamoun C., Gluzman I.Y., Goldberg D.E.;  
 RL "Plasmodium falciparum choline transporter (PfSCT1) gene."  
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.

Query Match 67.3%; Score 35; DB 5; Length 583;  
 Best Local Similarity 55.6%; Pred. No. 67;  
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
 QY 3 VVPXGMSYS 11  
 Db 227 IIPVGLSYS 235

RESULT 12  
 Q01487 PRELIMINARY; PRT; 670 AA.  
 AC Q01487;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
 DE DNA-binding protein AT-BP2 (ALPHA1-antitrypsin promoter binding protein 2) (Fragment).  
 DE Rattus norvegicus (Rat).  
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.  
 ON NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-SPRAGUE-DAWLEY; TISSUE=THYROID;  
 RX MEDLINE=91187610; PubMed=1901405;  
 RA Mitchelmore C., Traboni C., Cortese R.;  
 RT "Isolation of two cDNAs encoding zinc finger proteins which bind to the alpha 1-antitrypsin promoter and to the major histocompatibility complex class I enhancer."  
 RL Nucleic Acids Res. 19:141-147(1991).  
 CC -1- FUNCTION: BINDS TO THE ALPHA1-ANTITRYPSIN PROMOTER, TO THE KAPPA IMMUNOGLOBULIN GENE ENHANCER, AND TO THE MAJOR HISTOCOMPATIBILITY COMPLEX CLASS I ENHANCER: IT PLAYS A ROLE AS TRANSCRIPTIONAL REGULATOR. CONCERNING THE ALPHA1-ANTITRYPSIN EXPRESSION IT MAY ACT AS A NEGATIVE REGULATOR BY INTERFERING AT THE ALPHA1-ANTITRYPSIN PROMOTOR WITH ANOTHER NUCLEAR PROTEIN, CALLED LF-B.  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR.  
 CC -1- MISCELLANEOUS: ACIDIC AMINO ACID REGION MAY BE INVOLVED IN TRANSCRIPTIONAL ACTIVATION.  
 CC -1- SIMILARITY: HIGHLY SIMILAR TO AT-BP1 AND TO THE PRDII-BF1 FACTOR.  
 DR EMBL; X54250; CAA38151.1; -.  
 DR HSPSP; P15822; IBBO.  
 DR InterPro: IPR000822; Znf\_C2H2.  
 DR Pfam; PF00096; zf-C2H2; 2.  
 DR SMART; SM00355; Znf\_C2H2; 2.  
 DR PROSITE; PS00028; ZINC\_FINGER\_C2H2\_1; 2.  
 DR PROSITE; PS00157; ZINC\_FINGER\_C2H2\_2; 2.  
 KW Zinc-finger; Nuclear protein; DNA-binding; Transcription regulation;  
 KW Metal-binding; Multigene family.  
 FT NON\_TER 1  
 FT DOMAIN 54 104 ZINC-FINGERS.  
 FT DOMAIN 140 160 ACIDIC.  
 FT ZN\_FING 54 74 C(2)H(2) CLASS.  
 FT ZN\_FING 82 104 C(2)H(2) CLASS.  
 SQ SEQUENCE 670 AA; 71233 MW; CDD2324152590C17 CRC64;

Query Match 67.3%; Score 35; DB 11; Length 670;  
 Best Local Similarity 66.7%; Pred. No. 78;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 QY 3 VVPXGMSYS 11  
 Db 376 VVPAGLTYS 384

RESULT 13

DR EMBL; AY007374; AAK14818.1; -.  
 DR EMBL; AY007373; AAK14817.1; -.  
 DR InterPro: IPR002123; Acyltransferase.  
 DR Pfam; PF01553; Acyltransferase; 1.  
 SQ SEQUENCE 583 AA; 66918 MW; 4DF83D7530527474 CRC64;

Query Match 67.3%; Score 35; DB 5; Length 583;  
 Best Local Similarity 55.6%; Pred. No. 67;  
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMSYS 11  
 Db 227 IIPVGLSYS 235

RESULT 12

Q01487

PRELIMINARY; PRT; 670 AA.

AC Q01487;

DT 01-NOV-1996 (TREMBlrel. 01, Created)

DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)

DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)

DE DNA-binding protein AT-BP2 (ALPHA1-antitrypsin promoter binding protein 2) (Fragment).

DE Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.

ON NCBI\_TaxID=10116;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-SPRAGUE-DAWLEY; TISSUE=THYROID;

RX MEDLINE=91187610; PubMed=1901405;

RA Mitchelmore C., Traboni C., Cortese R.;

RT "Isolation of two cDNAs encoding zinc finger proteins which bind to the alpha 1-antitrypsin promoter and to the major histocompatibility complex class I enhancer."

RL Nucleic Acids Res. 19:141-147(1991).

CC -1- FUNCTION: BINDS TO THE ALPHA1-ANTITRYPSIN PROMOTER, TO THE KAPPA IMMUNOGLOBULIN GENE ENHANCER, AND TO THE MAJOR HISTOCOMPATIBILITY COMPLEX CLASS I ENHANCER: IT PLAYS A ROLE AS TRANSCRIPTIONAL REGULATOR. CONCERNING THE ALPHA1-ANTITRYPSIN EXPRESSION IT MAY ACT AS A NEGATIVE REGULATOR BY INTERFERING AT THE ALPHA1-ANTITRYPSIN PROMOTOR WITH ANOTHER NUCLEAR PROTEIN, CALLED LF-B.

CC -1- SUBCELLULAR LOCATION: NUCLEAR.

CC -1- MISCELLANEOUS: ACIDIC AMINO ACID REGION MAY BE INVOLVED IN TRANSCRIPTIONAL ACTIVATION.

CC -1- SIMILARITY: HIGHLY SIMILAR TO AT-BP1 AND TO THE PRDII-BF1 FACTOR.

DR EMBL; X54250; CAA38151.1; -.  
 DR HSPSP; P15822; IBBO.  
 DR InterPro: IPR000822; Znf\_C2H2.  
 DR Pfam; PF00096; zf-C2H2; 2.  
 DR SMART; SM00355; Znf\_C2H2; 2.  
 DR PROSITE; PS00028; ZINC\_FINGER\_C2H2\_1; 2.  
 DR PROSITE; PS00157; ZINC\_FINGER\_C2H2\_2; 2.  
 KW Zinc-finger; Nuclear protein; DNA-binding; Transcription regulation;  
 KW Metal-binding; Multigene family.  
 FT NON\_TER 1  
 FT DOMAIN 54 104 ZINC-FINGERS.  
 FT DOMAIN 140 160 ACIDIC.  
 FT ZN\_FING 54 74 C(2)H(2) CLASS.  
 FT ZN\_FING 82 104 C(2)H(2) CLASS.  
 SQ SEQUENCE 670 AA; 71233 MW; CDD2324152590C17 CRC64;

Query Match 67.3%; Score 35; DB 11; Length 670;  
 Best Local Similarity 66.7%; Pred. No. 78;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMSYS 11

Db 376 VVPAGLTYS 384

```

Q9PDM6
ID Q9PDM6 PRELIMINARY; PRT; 749 AA.
AC Q9PDM6;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Topoisomerase IV subunit.
GN XF1353.
OS Xylella fastidiosa.
OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
OC Xylella.
OC NCBI_TaxID=2371;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=9A5C;
RX MEDLINE=20365717; PubMed=10910347;
RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
RA Alvaranga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carrer H.,
RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorry H.,
RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
RA Fraga J.S., Franca S.C., Franco M.C., Frohne M., Furlan L.R.,
RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
RA Ho P.L., Hoheisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
RA Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
RA da Silveira J.F., Silvestri M.L.Z., Silveira W.J., de Souza A.A.,
RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tshun M.H.,
RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
RT "The genome sequence of the plant pathogen Xylella fastidiosa.";
RL Nature 406:151-159(2000).
DR EMBL; AF003967; AAF84162.1; -.
DR HSSP; P09097; 1AB4.
DR InterPro; IPR002205; DNA_topoisomIV.
DR Pfam; PF00521; DNA_topoisomIV; 1.
DR ProDom; PD000742; DNA_topoisomIV; 1.
DR SMART; SM00434; TOP4c; 1.
DR TIGRFAMs; TIGR01062; parC_Gneg; 1.
KW Complete proteome.
SQ SEQUENCE 749 AA; 83344 MW; 4DCD10F480EE0257 CRC64;

Query Match 67.3%; Score 35; DB 16; Length 749;
Best Local Similarity 77.8%; Pred. No. 88;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 EVVPXGMSY 10
DB 526 EVDPSGMSY 534

RESULT 14
Q14122
ID Q14122 PRELIMINARY; PRT; 1902 AA.
AC Q14122;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE DNA-binding protein (Mdp-21) (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90205817; PubMed=2108316;
RA Baldwin A.S., LeClair K.P., Singh H., Sharp P.A.;
RA "A large protein containing zinc finger domains binds to related
RT sequence elements in the enhancers of the class I major
RL histocompatibility complex and kappa immunoglobulin genes.";
RL Mol. Cell. Biol. 10:1406-1414(1990).
DR EMBL; M32019; AAA17534.1; -.
DR HSSP; P15822; 1BBO.
DR InterPro; IPR000822; Znf_C2H2.
DR Pfam; PF00096; zf-C2H2; 3.
DR SMART; SM00355; Znf_C2H2; 2.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 2.
DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 2.
DR DNA-binding; Metal-binding; Zinc-finger.
FT NON_TER 1
FT SEQUENCE 1902 AA; 207457 MW; 35BBF0D961E6ED20 CRC64;

Query Match 67.3%; Score 35; DB 4; Length 1902;
Best Local Similarity 66.7%; Pred. No. 2.4e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMSYS 11
DB 1590 VVPAGLTYS 1598

RESULT 15
Q12479
ID Q12479 PRELIMINARY; PRT; 156 AA.
AC Q12479;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE ORF YOR013W.
GN YOR013W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OC NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RA De haan M., Grivell L.A., Maarse A.C.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA MIPS;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA STRAIN=FY1679;
RC STRAIN=FY1679;
RA De haan M., Maarse A.C., Grivell L.A.;
RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=FY1679;
RX MEDLINE=94019318; PubMed=8413243;
RA Dumont M.E., Schlichter J.B., Cardillo T.S., Hayes M.K., Bethlenny G.,
RA Sherman F.;
RA "CYC2 encodes a factor involved in mitochondrial import of yeast
RT cytochrome C.";
RT Mol. Cell. Biol. 13:6442-6451(1993).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=FY1679;
RX MEDLINE=94169519; PubMed=7764548;
RA Lee Y.S., Shimizu J., Yoda K., Yanasaki M.;
RA "Molecular cloning of a gene, DHS1, which complements a drug-
RT hypersensitive mutation of the yeast Saccharomyces cerevisiae.";
RT Biosci. Biotechnol. Biochem. 58:391-395(1994).

```

```

DR EMBL; Z74920; CAA99201.1; -.
DR EMBL; X87331; CAA60762.1; -.
DR SGD; S0005539; YOR013W.
SQ SEQUENCE 156 AA; 17881 MW; 380442B74C272B41 CRC64;

Query Match      65.4%; Score 34; DB 3; Length 156;
Best Local Similarity 66.7%; Pred. No. 27;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 EVVPXGMSY 10
Db      ||:| |||
50 EVMPLGMDY 58

RESULT 16
Q971S2
ID Q971S2 PRELIMINARY; PRT; 219 AA.
AC Q971S2;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE putative ribose 5-phosphate isomerase.
GN ST1302.
OS Sulfolobus tokodaii.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.
OX NCBI_TaxID=111955;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JCM 10545 / 7;
RA PubMed=11572479;
RA Kwarababayasi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,
RA Sekine M., Baba S.-I., Ankai A., Kosugi H., Hosoyama A., Fukui S.,
RA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,
RA Yoshizawa T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,
RA Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A.,
RA Oshima T., Kikuchi H.;
RT "Complete genome sequence of an aerobic thermoacidophilic
RT Crenarchaeon, Sulfolobus tokodaii strain7."
RL DNA Res. 8:123-140(2001).
DR EMBL; AP000985; BAB66348.1; -.
DR InterPro; IPR004788; RpiA; 1.
DR ProDom; PD005813; RpiA; 1.
KW Isomerase; Hypothetical protein; Complete proteome.
SQ SEQUENCE 219 AA; 24541 MW; A4E9A3F2C4006D90 CRC64;

Query Match      65.4%; Score 34; DB 17; Length 219;
Best Local Similarity 66.7%; Pred. No. 38;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 EVVPXGMSY 10
Db      ||||| |::|
131 EVVPGVAV 139

RESULT 17
O28342
ID O28342 PRELIMINARY; PRT; 252 AA.
AC O28342;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Cell division inhibitor (MIND-2).
GN AF1937.
OS Archaeoglobus fulgidus.
OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
OC Archaeoglobaceae; Archaeoglobus.
OX NCBI_TaxID=2234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=9804943; PubMed=9389475;
RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,

```

```

RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
RA Richardson D.L., Kerlavage A.R., Graham D.E., Kyripides N.C.,
RA Fleischnmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Kirkness E.F., Dougherty B.A., Mckenney K., Adams M.D., Loftus B.,
RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
RA Cotton M.D., Spriggs T., Artiach P., Kaine B.P., Sykes S.M.,
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RT "The complete genome sequence of the hyperthermophilic, sulphate-
RT reducing archaeon Archaeoglobus fulgidus."
RL Nature 390:364-370(1997).
DR EMBL; AE000970; AAB89318.1; -.
DR TIGR; AF1937; -.
DR InterPro; IPR000707; ATPase_Para.
DR Pfam; PF00991; Para; 1.
KW Hypothetical protein; Cell division; Complete proteome.
SQ SEQUENCE 252 AA; 27130 MW; A401DC1F93E8C538 CRC64;

Query Match      65.4%; Score 34; DB 17; Length 252;
Best Local Similarity 75.0%; Pred. No. 44;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 EVVPXGMS 9
Db      ||:| |||
81 EVIPAGMS 88

RESULT 18
Q96MU1
ID Q96MU1 PRELIMINARY; PRT; 290 AA.
AC Q96MU1;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE CDNA FLJ31891 fis, clone NT2RP7003304, weakly similar to YceA protein
DE homolog ybfO.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Tashiro H., Yamazaki M., Watanabe K., Kumagai A., Itakura S.,
RA Fukuzumi Y., Fujimori Y., Komiyama M., Sugiyama T., Irie R.,
RA Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,
RA Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K., Wagatsuma M.,
RA Murakawa K., Kanehori K., Takahashi-Fujii A., Oshima A., Sugiyama A.,
RA Kawakami B., Suzuki Y., Sugano S., Nagahari K., Masuho Y., Nagai K.,
RA Isogai T.;
RT "NEDO human cDNA sequencing project."
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK056453; BAB71188.1; -.
DR InterPro; IPR001763; Rhodanese-like.
DR Pfam; PF00581; Rhodanese; 1.
SQ SEQUENCE 290 AA; 32972 MW; 41FB7FB0217C1421 CRC64;

Query Match      65.4%; Score 34; DB 4; Length 290;
Best Local Similarity 66.7%; Pred. No. 52;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXGMS 9
Db      ||:| |||
35 EEIVPMGIS 43

RESULT 19
Q92MD6
ID Q92MD6 PRELIMINARY; PRT; 387 AA.
AC Q92MD6;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)

```

DT	01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT	01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DT	Putative HIPPARTE hydrolase protein (EC 3.5.1.32).
GN	HIP01 OR R02690 OR SMC00682.
OS	Rhizobium melliloti (Sinorhizobium melliloti).
OC	Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC	Rhizobiaceae; Sinorhizobium.
NCBI_taxid=382;	
[1]	
RN	SEQUENCE FROM N.A.
RP	STRAIN=1021;
RC	MEDLINE=21396507; PubMed=11481430;
RA	Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,
RA	Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
RA	Godrie T., Coffeau A., Kahn D., Kiss E., Lelaure V., Masuy D.,
RA	Pohl T., Portetelle D., Puehler A., Purnelle B., Ramsperger U.,
RA	Renard C., Thebaud P., Vandenberg M., Weidner S., Galibert F.,
RT	"Analysis of the chromosome sequence of the legume symbiont
RT	Sinorhizobium melliloti strain 1021.";
RL	Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
DR	EMBL: AL591791; CAC47269.1; -
DR	InterPro: IPR002933; Peptidase.M20.
DR	Fram; PF01546; Peptide.M20; 1.
KW	Hydrolase; Complete proteome.
SQ	SEQUENCE 387 AA; 41074 MW; 342763088907A6E3 CRC64;
Query Match	65.4%; Score 34; DB 16; Length 387;
Best Local Similarity	50.0%; Pred. No. 70;
Matches	5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
Qy	1 EEVVPXGMSY 10 : : : : :
Dd	367 DEAIHPGISY 376
RESULT 20	
Q98BP5	PRELIMINARY; PRD; 541 AA.
ID	Q98BP5
AC	Q98BP5;
DC	01-OCT-2001 (TReMBLrel. 18, Created)
DT	01-OCT-2001 (TReMBLrel. 18, Last sequence update)
DT	01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE	Probable DNA ligase.
GN	MLJ5481.
OS	Rhizobium loti (Mesorhizobium loti).
OC	Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC	Phyllobacteriaceae; Mesorhizobium.
Ox	NCBI_taxid=381;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=MAFF303099;
RX	MEDLINE=21082930; PubMed=11214968;
RA	Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA	Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA	Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA	Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RA	Takeuchi C., Yamada M., Tabata S.
RT	"Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT	Mesorhizobium loti.";
RL	DNA Res. 7:331-338(2000).
DR	EMBL: AF003006; BAB51927.1; -
DR	InterPro: IPR000977; DNA_ligase.
DR	Pfam: PF01068; DNA_ligase; 1.
DR	PROSITE: PS00697; DNA_LIGASE_1; UNKNOWN_1.
DR	PROSITE: PS0160; DNA_LIGASE_A3; 1.
KW	Ligase; Complete proteome.
SQ	SEQUENCE 541 AA; 60645 MW; 2EFEF70543F28F8 CRC64;
Query Match	65.4%; Score 34; DB 16; Length 541;
Best Local Similarity	60.0%; Pred. No. 1e+02;
Matches	6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
Qy	1 EEVVPXGMSY 10

```

QY 3 VVPXGMSYS 11
Db 111 VVPOGMAYA 119

RESULT 23
Q8XT05 PRELIMINARY; PRT; 1049 AA.
AC Q8XT05;
DT 01-MAR-2002 (TremBLrel. 20, Created)
DT 01-MAR-2002 (TremBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TremBLrel. 21, Last annotation update)
DE Probable multidrug efflux system transmembrane protein.
GN MEXD OR RSP0312 OR RS05457.
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OG Plasmid megaplasmid.
OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
OC Ralstonia.
OX NCBI_TaxID=305;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=GM11000;
RX MEDLINE=21681879; PubMed=11823852;
RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Manganot S.,
RA Arlat M., Billault A., Brottier P., Camus J.C., Cattolico L.,
RA Chandlin M., Choise N., Claudel-Renard C., Cunnac S., Demange N.,
RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schiex T.,
RA Sigulier P., Thebaud P., Whalen M., Wincker P., Levy M.,
RA Wellesbach J., Boucher C.A.;
RT Genome sequence of the plant pathogen Ralstonia solanacearum.;
RL Nature 415:497-502(2002).
DR EMBL; AL646078; CAB17463.1; -
DR InterPro; IPR001036; Acrflavin_res.
DR InterPro; IPR004764; HAE1.
DR InterPro; IPR000731; HMGCR/patch_5TM.
DR Pfam; PF00873; ACR_tran; 1.
DR PRINTS; PR00702; ACRIFLAVINRP.
DR TIGRFAMS; TIGR00915; 2A0602; 1.
DR PROSITE; PS50156; SSD; 1.
KW Plasmid; Complete proteome.
SQ SEQUENCE 1049 AA; 111769 MW; CB59674B670089CE CRC64;

Query Match 65.4%; Score 34; DB 16; Length 1049;
Best Local Similarity 75.0%; Pred. No. 2.1e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 VVPXGMSYS 11
Db 317 MPAGMSYS 324

RESULT 24
Q96914 PRELIMINARY; PRT; 1499 AA.
AC Q96914;
DT 01-DEC-2001 (TremBLrel. 19, Created)
DT 01-DEC-2001 (TremBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TremBLrel. 21, Last annotation update)
DE Putative aminophospholipid translocase (Aminophospholipid-transporting
DE ATPase).
GN ATP10C.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21225279; PubMed=11326269;
RA Meguro M., Kashiwagi A., Mitsuya K., Nakao M., Kondo I., Saitoh S.,
RA Oshimura M.;
RT "A novel maternally expressed gene, ATP10C, encodes a putative
RT aminophospholipid translocase associated with Angelman syndrome.";

```

```

RL Nat. Genet. 28:19-20(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21313119; PubMed=11353404;
RA Herzig L.B.K., Kim S.J., Cook E.H. Jr., Ledbetter D.H.;
RT "The human aminophospholipid-transporting ATPase gene ATP10C maps
RT adjacent to UBE3A and exhibits similar imprinted expression.";
RL Am. J. Hum. Genet. 68:1501-1505(2001).
DR EMBL; AB051358; BAB47392.1; -
DR EMBL; AY029504; AAK33100.1; JOINED.
DR EMBL; AY029487; AAK33100.1; JOINED.
DR EMBL; AY029488; AAK33100.1; JOINED.
DR EMBL; AY029489; AAK33100.1; JOINED.
DR EMBL; AY029490; AAK33100.1; JOINED.
DR EMBL; AY029491; AAK33100.1; JOINED.
DR EMBL; AY029492; AAK33100.1; JOINED.
DR EMBL; AY029493; AAK33100.1; JOINED.
DR EMBL; AY029494; AAK33100.1; JOINED.
DR EMBL; AY029495; AAK33100.1; JOINED.
DR EMBL; AY029496; AAK33100.1; JOINED.
DR EMBL; AY029497; AAK33100.1; JOINED.
DR EMBL; AY029498; AAK33100.1; JOINED.
DR EMBL; AY029499; AAK33100.1; JOINED.
DR EMBL; AY029500; AAK33100.1; JOINED.
DR EMBL; AY029501; AAK33100.1; JOINED.
DR EMBL; AY029502; AAK33100.1; JOINED.
DR EMBL; AY029503; AAK33100.1; JOINED.
DR InterPro; IPR001757; ATPase_E1-E2.
DR InterPro; IPR001054; Hlgnae/hydrolase.
DR InterPro; IPR001454; Hydrolase; 1.
DR Pfam; PF00702; Hydrolase; 1.
DR PROSITE; PS00154; ATPASE_E1_E2; UNKNOWN_1.
DR PROSITE; PS00225; CRYSTALLIN_BETAGAMMA; UNKNOWN_1.
SQ SEQUENCE 1499 AA; 167687 MW; D4996A4D0635A68D CRC64;

Query Match 65.4%; Score 34; DB 4; Length 1499;
Best Local Similarity 72.7%; Pred. No. 3e+02;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMSYS 11
Db 469 EEVVPXGMSYS 479

RESULT 25
Q8TX62 PRELIMINARY; PRT; 143 AA.
AC Q8TX62;
DT 01-JUN-2002 (TremBLrel. 21, Created)
DT 01-JUN-2002 (TremBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TremBLrel. 21, Last annotation update)
DE Uncharacterized conserved protein.
GN MK0814.
OS Methanopyrus kandleri.
OC Archaea; Euryarchaeota; Methanopyri; Methanopyrales; Methanopyraceae;
OC Methanopyrus.
OX NCBI_TaxID=2320;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=AV19 / DSM 6324 / JCM 9639;
RX MEDLINE=21927647; PubMed=11930014;
RA Slesarev A.I., Mezhevaya K.V., Makarova K.S., Polushin N.N.,
RA Scherbinina O.V., Shakhova V.V., Belova G.I., Aravind L.,
RA Natale D.A., Koonin I.B., Tatusov R.L., Wolf Y.I., Stetter K.O.,
RA Malykh A.G., Koonin E.V., Kozlyavkin S.A.;
RT "The complete genome of hyperthermophile Methanopyrus kandleri AV19
RT and monophyly of archaeal methanogens."
RL Proc. Natl. Acad. Sci. U.S.A. 99:4644-4649(2002).
DR EMBL; AE010372; AAK02027.1; -
DR Complete proteome.
KW Complete proteome.
SQ SEQUENCE 143 AA; 15734 MW; 4C8B28A1FBBEDD0B CRC64;

Query Match 63.5%; Score 33; DB 17; Length 143;

```



Wed Jun 11 15:42:06 2003

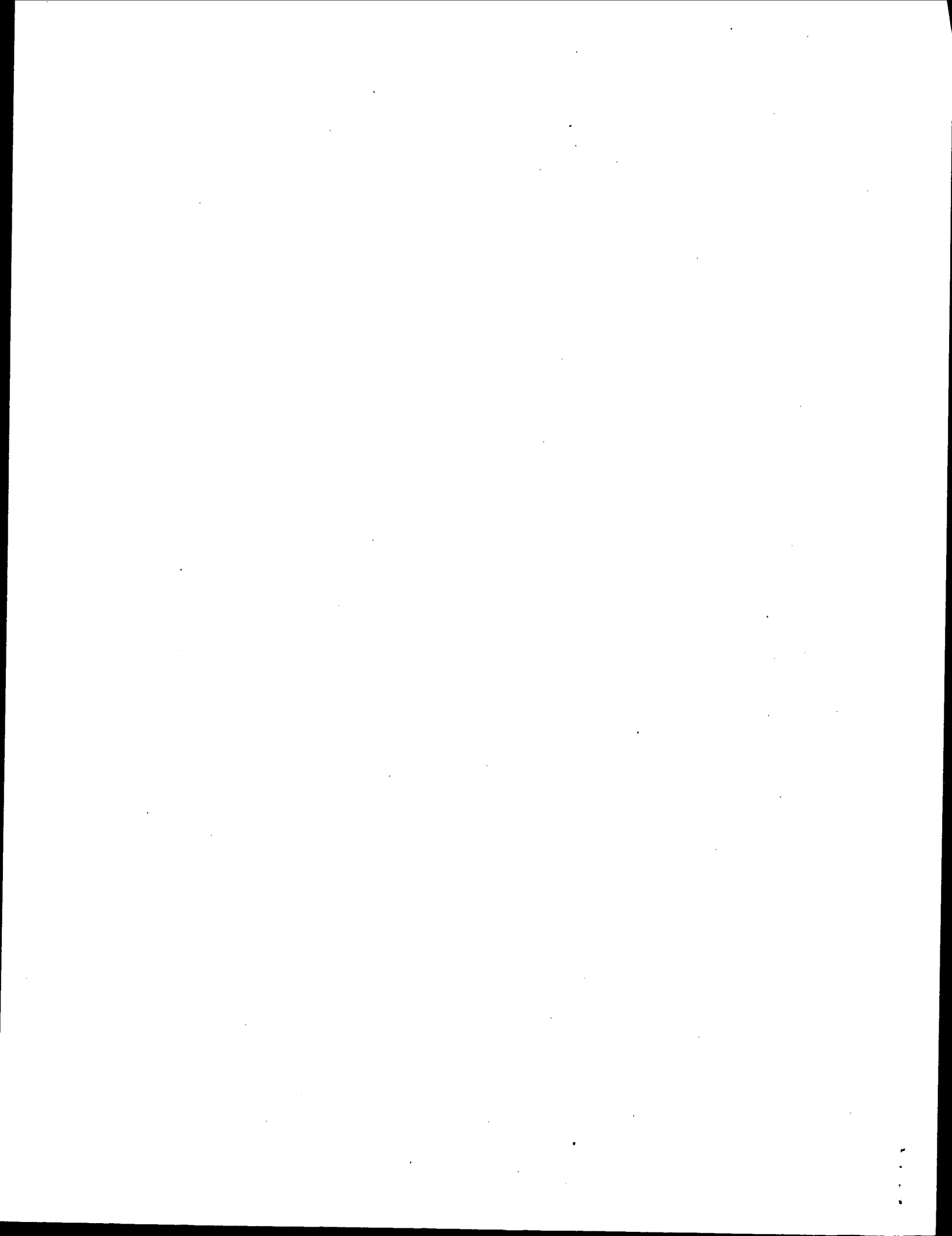
us-09-909-164-10.rspt

Best Local Similarity 60.0%; Pred. No. 40;  
Matches 6; Conservative 1; Mismatches

3; Indels 0; Gaps 0;

Qy 1 EEVVPXGMSY 10  
||:| |  
Db 75 EELVPQGAGY 84

Search completed: June 10, 2003, 13:46:32  
Job time : 25.7857 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2003, 13:24:45 ; Search time 31.3571 Seconds  
(without alignments)  
46.744 Million cell updates/sec

Title: US-09-909-164-11  
Perfect score: 56  
Sequence: 1 EEVVPXGMHYS 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 75 summaries

Database : A\_Geneseq\_101002.\*  
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	54	96.4	11	23	ABB80523 Hepatitis C virus
2	54	96.4	11	23	ABB80527 Hepatitis C virus
3	54	96.4	11	23	ABB80558 Hepatitis C virus
4	54	96.4	11	23	ABB80560 Hepatitis C virus
5	49	87.5	11	23	ABB80537 Hepatitis C virus
6	49	87.5	11	23	ABB80541 Hepatitis C virus
7	48	85.7	11	23	ABB80546 Hepatitis C virus
8	48	85.7	11	23	ABB80550 Hepatitis C virus
9	48	85.7	11	23	ABB80554 Hepatitis C virus
10	48	85.7	11	23	ABB80555 Hepatitis C virus

11	46	82.1	11	23	ABB80531 Hepatitis C virus
12	46	82.1	11	23	ABB80532 Hepatitis C virus
13	45	80.4	11	23	ABB80521 Hepatitis C virus
14	45	80.4	11	23	ABB80522 Hepatitis C virus
15	45	80.4	11	23	ABB80524 Hepatitis C virus
16	45	80.4	11	23	ABB80525 Hepatitis C virus
17	45	80.4	11	23	ABB80526 Hepatitis C virus
18	45	80.4	11	23	ABB80528 Hepatitis C virus
19	45	80.4	11	23	ABB80529 Hepatitis C virus
20	45	80.4	11	23	ABB80539 Hepatitis C virus
21	45	80.4	11	23	ABB80561 Hepatitis C virus
22	45	80.4	11	23	ABB80562 Hepatitis C virus
23	45	80.4	11	23	ABB80563 Hepatitis C virus
24	45	80.4	11	23	ABB80564 Hepatitis C virus
25	45	80.4	11	23	ABB80565 Hepatitis C virus
26	45	80.4	11	23	ABB80566 Hepatitis C virus
27	45	80.4	11	23	ABB80567 Hepatitis C virus
28	45	80.4	11	23	ABB80568 Hepatitis C virus
29	40	71.4	11	23	ABB80535 Hepatitis C virus
30	40	71.4	11	23	ABB80536 Hepatitis C virus
31	40	71.4	11	23	ABB80538 Hepatitis C virus
32	40	71.4	11	23	ABB80539 Hepatitis C virus
33	40	71.4	11	23	ABB80540 Hepatitis C virus
34	40	71.4	11	23	ABB80542 Hepatitis C virus
35	40	71.4	11	23	ABB80543 Hepatitis C virus
36	39	69.6	11	23	ABB80544 Hepatitis C virus
37	39	69.6	11	23	ABB80545 Hepatitis C virus
38	39	69.6	11	23	ABB80547 Hepatitis C virus
39	39	69.6	11	23	ABB80548 Hepatitis C virus
40	39	69.6	11	23	ABB80549 Hepatitis C virus
41	39	69.6	11	23	ABB80551 Hepatitis C virus
42	39	69.6	11	23	ABB80552 Hepatitis C virus
43	39	69.6	11	23	ABB80553 Hepatitis C virus
44	39	69.6	11	23	ABB80556 Hepatitis C virus
45	39	69.6	11	23	ABB80557 Hepatitis C virus
46	39	69.6	161	21	AA1940435 Human ORF199
47	39	69.6	161	21	ABP07096 Human ORF199
48	39	69.6	567	22	AAU39521 Propionibacterium
49	39	69.6	1037	23	ABP39949 Staphylococcus epi
50	37	66.1	11	23	ABB80530 Hepatitis C virus
51	37	66.1	11	23	ABB80533 Hepatitis C virus
52	37	66.1	11	23	ABB80534 Hepatitis C virus
53	36	64.3	55	23	AAW50964 Hepatitis C virus
54	36	64.3	743	21	AAV99488 Yeast acyltransfer
55	35	62.5	20	20	AAU76810 Hepatitis C virus
56	35	62.5	1022	22	ABG03621 Novel human diagno
57	35	62.5	1022	22	ABG05826 Novel human diagno
58	35	62.5	1022	22	ABG08173 Novel human diagno
59	34	60.7	150	22	AAU50013 Propionibacterium
60	34	60.7	637	22	ABBS7826 Drosophila melanog
61	34	60.7	2438	22	ABBS9970 Novel human diagno
62	33	58.9	80	22	ABG21296 Human immune/haema
63	33	58.9	92	22	AAW66603 Human secreted pro
64	33	58.9	200	23	AAU91410 Human albumin fusi
65	33	58.9	231	23	ABG65016 Human secreted pro
66	33	58.9	231	23	AAU91390 Human albumin fusi
67	33	58.9	240	23	ABG65018 Human secreted pro
68	33	58.9	240	23	AAU91360 Human polypeptide
69	33	58.9	257	22	AAO10294 Human secreted pro
70	33	58.9	286	23	AAU91409 Murine soluble int
71	33	58.9	379	21	AAV59390 Novel human diagno
72	33	58.9	394	22	ABG07474 Human endocrine po
73	33	58.9	445	22	AAU18340 Novel human diagno
74	33	58.9	449	22	ABG07475 Amino acid sequenc
75	33	58.9	465	20	AAV36985

## ALIGNMENTS

RESULT 1  
ABB80523  
ID ABB80523 standard; peptide: 11 AA.

XX AC ABB80523;  
 XX DT 08-OCT-2002 (first entry)  
 XX DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #3.  
 XX KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
 XX KW virucide.  
 XX OS Synthetic.  
 XX FH Key Location/Qualifiers  
 XX FT Modified-site 1  
 XX FT Modified-site 6 /note= "N-terminal acetyl"  
 XX FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with  
 XX FT residue 7"  
 XX OS Synthetic.  
 XX FH Key Location/Qualifiers  
 XX FT Modified-site 1 /note= "N-terminal acetyl"  
 XX FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with  
 XX FT residue 7"  
 XX FT Misc-difference 9  
 XX FT Modified-site 11 /note= "D-form residue"  
 XX FT Modified-site 11 /note= "C-terminal amide"  
 XX PN WO200208251-A2.  
 XX PD 31-JAN-2002.  
 XX PF 19-JUL-2001; 2001WO-US23169.  
 XX PR 21-JUL-2000; 2000US-220101P.  
 XX PA (CORV-) CORVAS INT INC.  
 XX PI Lim-wilby M, Levy OE, Brunck TK;  
 XX XX WPI; 2002-361643/39.  
 XX DR Novel peptide compound having hepatitis C virus protease inhibitory  
 XX PT activity useful for treating disorders associated with hepatitis C  
 XX PT virus protease -  
 XX PS Claim 17; Page 64; 69pp; English.  
 XX CC The sequence represents a peptide compound of the invention having  
 XX CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
 XX CC invention are alpha-ketoamide peptide analogues. The peptides have  
 XX CC virucide activity, and are useful for treating and in the manufacture of  
 XX CC a medicament to treat disorders associated with HCV protease. A  
 XX CC pharmaceutical composition comprising the peptide as an active ingredient  
 XX CC is useful for treating disorders associated with hepatitis C virus.  
 XX SQ Sequence 11 AA;  
 XX Query Match 96.4%; Score 54; DB 23; Length 11;  
 XX Best Local Similarity 100.0%; Pred. No. 0.00045;  
 XX Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 EEVVPXGMHYS 11  
 DB 1 EEVVPXGMHYS 11  
 RESULT 2  
 ABB80527  
 ID ABB80527 standard; peptide; 11 AA.  
 XX AC ABB80527;  
 XX DT 08-OCT-2002 (first entry)  
 XX DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #7.  
 XX KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;

KW virucide.  
 XX OS Synthetic.  
 XX FH Key Location/Qualifiers  
 XX FT Modified-site 1  
 XX FT Modified-site 6 /note= "N-terminal acetyl"  
 XX FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with  
 XX FT residue 7"  
 XX FT Misc-difference 8  
 XX FT Modified-site 11 /note= "D-form residue"  
 XX FT Modified-site 11 /note= "C-terminal amide"  
 XX PN WO200208251-A2.  
 XX PD 31-JAN-2002.  
 XX PF 19-JUL-2001; 2001WO-US23169.  
 XX PR 21-JUL-2000; 2000US-220101P.  
 XX PA (CORV-) CORVAS INT INC.  
 XX PI Lim-wilby M, Levy OE, Brunck TK;  
 XX XX WPI; 2002-361643/39.  
 XX DR Novel peptide compound having hepatitis C virus protease inhibitory  
 XX PT activity useful for treating disorders associated with hepatitis C  
 XX PT virus protease -  
 XX PS Claim 17; Page 64; 69pp; English.  
 XX CC The sequence represents a peptide compound of the invention having  
 XX CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
 XX CC invention are alpha-ketoamide peptide analogues. The peptides have  
 XX CC virucide activity, and are useful for treating and in the manufacture of  
 XX CC a medicament to treat disorders associated with HCV protease. A  
 XX CC pharmaceutical composition comprising the peptide as an active ingredient  
 XX CC is useful for treating disorders associated with hepatitis C virus.  
 XX SQ Sequence 11 AA;  
 XX Query Match 96.4%; Score 54; DB 23; Length 11;  
 XX Best Local Similarity 100.0%; Pred. No. 0.00045;  
 XX Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 EEVVPXGMHYS 11  
 DB 1 EEVVPXGMHYS 11  
 RESULT 3  
 ABB80558  
 ID ABB80558 standard; peptide; 11 AA.  
 XX AC ABB80558;  
 XX DT 08-OCT-2002 (first entry)  
 XX DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #38.  
 XX KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
 XX KW virucide.  
 XX OS Synthetic.  
 XX FH Key Location/Qualifiers  
 XX FT Modified-site 1  
 XX FT Modified-site 6 /note= "N-terminal acetyl";

FT FT /note= "Norvalyl carbonyl forming keto-amide linkage with  
FT residue 7"  
FT Modified-site 8  
FT PN /note= "Oxymethionine"  
FT Modified-site 11  
FT FT /note= "C-terminal amide"  
XX XX  
XX PN WO200208251-A2.  
XX PD 31-JAN-2002.  
XX XX  
XX PF 19-JUL-2001; 2001WO-US23169.  
XX PR 21-JUL-2000; 2000US-220101P.  
XX PA (CORV-) CORVAS INT INC.  
XX PI Lim-wilby M, Levy OE, Brunck TK;  
XX DR WPI; 2002-361643/39.  
XX Novel peptide compound having hepatitis C virus protease inhibitory  
PT activity useful for treating disorders associated with hepatitis C  
PT virus protease  
XX Claim 17; Page 65; 69pp; English.  
XX The sequence represents a peptide compound of the invention having  
CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
CC invention are alpha-ketoamide peptide analogues. The peptides have  
CC virucide activity, and are useful for treating and in the manufacture of  
CC a medicament to treat disorders associated with HCV protease. A  
CC pharmaceutical composition comprising the peptide as an active ingredient  
CC is useful for treating disorders associated with hepatitis C virus.  
XX Sequence 11 AA;  
SQ

Query Match 96.4%; Score 54; DB 23; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.00045;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEVVPXGMHYS 11  
DB 1 EEVVPXGMHYS 11  
IIIIIIIIII

RESULT 4  
ABB80560  
ID ABB80560 standard; peptide; 11 AA.  
XX AC ABB80560;  
XX XX  
XX 08-OCT-2002 (first entry)  
XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #40.  
XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
KW virucide.  
XX Synthetic.  
XX OS Location/Qualifiers  
FH Key  
FT Modified-site 1 /note= "N-terminal acetyl"  
FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with  
FT residue 7"  
FT Misc-difference 8 /note= "D-form residue"  
FT Modified-site 8 /note= "Oxymethionine"  
FT Misc-difference 9 /note= "D-form residue"  
FT

FT Modified-site 11 /note= "C-terminal amide"  
FT  
XX XX WO200208251-A2.  
XX PN 31-JAN-2002.  
XX PD  
XX PF 19-JUL-2001; 2001WO-US23169.  
XX PR 21-JUL-2000; 2000US-220101P.  
XX PA (CORV-) CORVAS INT INC.  
XX PI Lim-wilby M, Levy OE, Brunck TK;  
XX DR WPI; 2002-361643/39.  
XX Novel peptide compound having hepatitis C virus protease inhibitory  
PT activity useful for treating disorders associated with hepatitis C  
PT virus protease  
XX Claim 17; Page 65; 69pp; English.  
XX The sequence represents a peptide compound of the invention having  
CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
CC invention are alpha-ketoamide peptide analogues. The peptides have  
CC virucide activity, and are useful for treating and in the manufacture of  
CC a medicament to treat disorders associated with HCV protease. A  
CC pharmaceutical composition comprising the peptide as an active ingredient  
CC is useful for treating disorders associated with hepatitis C virus.  
XX Sequence 11 AA;  
SQ

Query Match 96.4%; Score 54; DB 23; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.00045;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EEVVPXGMHYS 11  
DB 1 EEVVPXGMHYS 11  
IIIIIIIIII

RESULT 5  
ABB80537  
ID ABB80537 standard; peptide; 11 AA.  
XX AC ABB80537;  
XX XX  
XX 08-OCT-2002 (first entry)  
XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #17.  
XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
KW virucide.  
XX Synthetic.  
XX OS Location/Qualifiers  
FH Key  
FT Modified-site 1 /note= "N-terminal acetyl"  
FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with  
FT residue 7"  
FT Misc-difference 9 /note= "D-form residue"  
FT Modified-site 11 /note= "C-terminal amide"  
FT  
XX XX WO200208251-A2.  
XX PN 31-JAN-2002.  
XX PD  
XX PF 19-JUL-2001; 2001WO-US23169.

XX PR 21-JUL-2000; 2000US-220101P.  
 XX PA (CORV-) CORVAS INT INC.  
 XX PI Lim-wilby M, Levy OE, Brunk TK;  
 XX PS WPI; 2002-361643/39.  
 XX DR  
 XX PT Novel peptide compound having hepatitis C virus protease inhibitory  
 PT activity useful for treating disorders associated with hepatitis C  
 PT virus protease -  
 XX CC  
 XX PS Claim 17; Page 64; 69pp; English.  
 XX CC  
 XX CC The sequence represents a peptide compound of the invention having  
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
 CC invention are alpha-ketoamide peptide analogues. The peptides have  
 CC virucide activity, and are useful for treating and in the manufacture of  
 CC a medicament to treat disorders associated with HCV protease. A  
 CC pharmaceutical composition comprising the peptide as an active ingredient  
 CC is useful for treating disorders associated with hepatitis C virus.  
 XX CC  
 XX SQ Sequence 11 AA;  
 Query Match 87.5%; Score 49; DB 23; Length 11;  
 Best Local Similarity 90.9%; Pred. No. 0.0043;  
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 EEVVPXGMHYS 11  
 DB 1 EEVVPXGMHYS 11  
 RESULT 6  
 ABB80541  
 ID ABB80541 standard; peptide; 11 AA.  
 XX AC ABB80541;  
 XX XX  
 XX DT 08-OCT-2002 (first entry)  
 XX DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #21.  
 XX KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
 KW virucide.  
 XX OS Synthetic.  
 XX XX  
 XX FH Key Location/Qualifiers  
 FT Modified-site 1 /note= "N-terminal acetyl"  
 FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with  
 FT residue 7"  
 FT Misc-difference 8 /note= "D-form residue"  
 FT Modified-site 11 /note= "C-terminal amide"  
 FT XX  
 XX PN W0200208251-A2.  
 XX PD 31-JAN-2002.  
 XX PF 19-JUL-2001; 2001WO-US23169.  
 XX PR 21-JUL-2000; 2000US-220101P.  
 XX XX  
 XX PA (CORV-) CORVAS INT INC.  
 XX XX  
 XX PI Lim-wilby M, Levy OE, Brunk TK;  
 XX XX  
 XX DR WPI; 2002-361643/39.

XX PT Novel peptide compound having hepatitis C virus protease inhibitory  
 PT activity useful for treating disorders associated with hepatitis C  
 PT virus protease -  
 XX CC  
 XX PS Claim 17; Page 65; 69pp; English.  
 XX CC  
 XX CC The sequence represents a peptide compound of the invention having  
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
 CC invention are alpha-ketoamide peptide analogues. The peptides have  
 CC virucide activity, and are useful for treating and in the manufacture of  
 CC a medicament to treat disorders associated with HCV protease. A  
 CC pharmaceutical composition comprising the peptide as an active ingredient  
 CC is useful for treating disorders associated with hepatitis C virus.  
 XX CC  
 XX SQ Sequence 11 AA;  
 Query Match 87.5%; Score 49; DB 23; Length 11;  
 Best Local Similarity 90.9%; Pred. No. 0.0043;  
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 EEVVPXGMHYS 11  
 DB 1 EEVVPXGMHYS 11  
 RESULT 7  
 ABB80546  
 ID ABB80546 standard; peptide; 11 AA.  
 XX AC ABB80546;  
 XX XX  
 XX DT 08-OCT-2002 (first entry)  
 XX DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #26.  
 XX KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
 KW virucide.  
 XX OS Synthetic.  
 XX XX  
 XX FH Key Location/Qualifiers  
 FT Modified-site 1 /note= "N-terminal acetyl"  
 FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with  
 FT residue 7"  
 FT Modified-site 11 /note= "C-terminal amide"  
 FT XX  
 XX PN W0200208251-A2.  
 XX PD 31-JAN-2002.  
 XX PF 19-JUL-2001; 2001WO-US23169.  
 XX PR 21-JUL-2000; 2000US-220101P.  
 XX XX  
 XX PA (CORV-) CORVAS INT INC.  
 XX XX  
 XX PI Lim-wilby M, Levy OE, Brunk TK;  
 XX XX  
 XX DR WPI; 2002-361643/39.  
 XX XX  
 XX PT Novel peptide compound having hepatitis C virus protease inhibitory  
 PT activity useful for treating disorders associated with hepatitis C  
 PT virus protease -  
 XX CC  
 XX PS Claim 17; Page 65; 69pp; English.  
 XX CC  
 XX CC The sequence represents a peptide compound of the invention having  
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
 CC invention are alpha-ketoamide peptide analogues. The peptides have  
 CC virucide activity, and are useful for treating and in the manufacture of  
 CC a medicament to treat disorders associated with HCV protease. A  
 CC pharmaceutical composition comprising the peptide as an active ingredient  
 CC is useful for treating disorders associated with hepatitis C virus.  
 XX CC

virucide activity, and are useful for treating and in the manufacture of a medicament to treat disorders associated with HCV protease. A pharmaceutical composition comprising the peptide as an active ingredient is useful for treating disorders associated with hepatitis C virus.

Sequence 11 AA;

Query Match 85.7%; Score 48; DB 23; Length 11;

Best Local Similarity 90.9%; Pred. No. 0.0067; Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXGMHYS 11  
 Db 1 EEVVPXGMHYS 11

# RESULT 8

ABB80550  
 ID ABB80550 standard; peptide; 11 AA.

XX AC ABB80550;

DT 08-OCT-2002 (first entry)

XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #30.

XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide; virucide.

XX Synthetic.

XX Key Location/Qualifiers

FT Modified-site 1 /note= "N-terminal acetyl"

FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with residue 7"

FT Misc-difference 9 /note= "D-form residue"

FT Modified-site 11 /note= "C-terminal amide"

XX WO200208251-A2.

XX 31-JAN-2002.

XX 19-JUL-2001; 2001WO-US23169.

XX 21-JUL-2000; 2000US-220101P.

XX (CORV-) CORVAS INT INC.

XX Lim-wilby M, Levy OE, Brunc TK;

XX WPI; 2002-361643/39.

XX Novel peptide compound having hepatitis C virus protease inhibitory activity useful for treating disorders associated with hepatitis C virus protease

XX Claim 17; Page 65; 69pp; English.

XX The sequence represents a peptide compound of the invention having hepatitis C virus (HCV) protease inhibitory activity. The peptides of the invention are alpha-ketoamide peptide analogues. The peptides have virucide activity, and are useful for treating and in the manufacture of a medicament to treat disorders associated with HCV protease. A pharmaceutical composition comprising the peptide as an active ingredient is useful for treating disorders associated with hepatitis C virus.

XX Query Match 85.7%; Score 48; DB 23; Length 11;

XX Best Local Similarity 90.9%; Pred. No. 0.0067; Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXGMHYS 11

Db 1 EEVVPXGMHYS 11

Best Local Similarity 90.9%; Pred. No. 0.0067; Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXGMHYS 11  
 Db 1 EEVVPXGMHYS 11

# RESULT 9

ABB80554  
 ID ABB80554 standard; peptide; 11 AA.

XX AC ABB80554;

DT 08-OCT-2002 (first entry)

XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #34.

XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide; virucide.

XX Synthetic.

XX Key Location/Qualifiers

FT Modified-site 1 /note= "N-terminal acetyl"

FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with residue 7"

FT Misc-difference 8 /note= "D-form residue"

FT Modified-site 11 /note= "C-terminal amide"

XX WO200208251-A2.

XX 31-JAN-2002.

XX 19-JUL-2001; 2001WO-US23169.

XX 21-JUL-2000; 2000US-220101P.

XX (CORV-) CORVAS INT INC.

XX Lim-wilby M, Levy OE, Brunc TK;

XX WPI; 2002-361643/39.

XX Novel peptide compound having hepatitis C virus protease inhibitory activity useful for treating disorders associated with hepatitis C virus protease

XX Claim 17; Page 65; 69pp; English.

XX The sequence represents a peptide compound of the invention having hepatitis C virus (HCV) protease inhibitory activity. The peptides of the invention are alpha-ketoamide peptide analogues. The peptides have virucide activity, and are useful for treating and in the manufacture of a medicament to treat disorders associated with HCV protease. A pharmaceutical composition comprising the peptide as an active ingredient is useful for treating disorders associated with hepatitis C virus.

XX Query Match 85.7%; Score 48; DB 23; Length 11;

XX Best Local Similarity 90.9%; Pred. No. 0.0067; Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXGMHYS 11

Db 1 EEVVPXGMHYS 11

RESULT 10  
 ABB80555  
 ID ABB80555 standard; peptide; 11 AA.  
 AC ABB80555;  
 XX  
 DT 08-OCT-2002 (first entry)  
 XX  
 DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #35.  
 XX  
 KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
 KW virucide.  
 XX  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Modified-site 1 /note= "N-terminal acetyl"  
 FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with  
 FT residue 7"  
 FT Misc-difference 8 /note= "D-form residue"  
 FT Misc-difference 9 /note= "D-form residue"  
 FT Modified-site 11 /note= "C-terminal amide"  
 FT  
 PN WO200208251-A2.  
 XX  
 XX 31-JAN-2002.  
 PD  
 PF 19-JUL-2001; 2001WO-US23169.  
 XX  
 PR 21-JUL-2000; 2000US-220101P.  
 XX  
 PA (CORV-) CORVAS INT INC.  
 XX  
 PI Lim-wilby M, Levy OE, Brunck TK;  
 XX  
 PI WPI; 2002-361643/39.  
 DR  
 XX Novel peptide compound having hepatitis C virus protease inhibitory  
 PT activity useful for treating disorders associated with hepatitis C  
 PT virus protease  
 XX  
 PS Claim 17; Page 65; 69pp; English.  
 XX  
 CC The sequence represents a peptide compound of the invention having  
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
 CC invention are alpha-ketoamide peptide analogues. The peptides have  
 CC virucide activity, and are useful for treating and in the manufacture of  
 CC a medicament to treat disorders associated with HCV protease. A  
 CC pharmaceutical composition comprising the peptide as an active ingredient  
 CC is useful for treating disorders associated with hepatitis C virus.  
 XX  
 SQ Sequence 11 AA;  
 Query Match 85.7%; Score 48; DB 23; Length 11;  
 Best Local Similarity 90.9%; Pred. No. 0.0067;  
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 EEVVPXGMHYS 11  
 Db  
 1 EEVVPXGSHYS 11  
 XX  
 RESULT 11  
 ABB80531  
 ID ABB80531 standard; peptide; 11 AA.  
 AC ABB80531;  
 XX

DT 08-OCT-2002 (first entry)  
 XX  
 DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #11.  
 XX  
 KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
 KW virucide.  
 XX  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Modified-site 1 /note= "N-terminal acetyl"  
 FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with  
 FT residue 7"  
 FT Modified-site 11 /note= "C-terminal amide"  
 FT  
 PN WO200208251-A2.  
 XX  
 XX 31-JAN-2002.  
 PD  
 PF 19-JUL-2001; 2001WO-US23169.  
 XX  
 PR 21-JUL-2000; 2000US-220101P.  
 XX  
 PA (CORV-) CORVAS INT INC.  
 XX  
 PI Lim-wilby M, Levy OE, Brunck TK;  
 XX  
 PI WPI; 2002-361643/39.  
 DR  
 XX Novel peptide compound having hepatitis C virus protease inhibitory  
 PT activity useful for treating disorders associated with hepatitis C  
 PT virus protease  
 XX  
 PS Claim 17; Page 64; 69pp; English.  
 XX  
 CC The sequence represents a peptide compound of the invention having  
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
 CC invention are alpha-ketoamide peptide analogues. The peptides have  
 CC virucide activity, and are useful for treating and in the manufacture of  
 CC a medicament to treat disorders associated with HCV protease. A  
 CC pharmaceutical composition comprising the peptide as an active ingredient  
 CC is useful for treating disorders associated with hepatitis C virus.  
 XX  
 SQ Sequence 11 AA;  
 Query Match 82.1%; Score 46; DB 23; Length 11;  
 Best Local Similarity 90.9%; Pred. No. 0.017;  
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 EEVVPXGMHYS 11  
 Db  
 1 EEVVPXGSHYS 11  
 XX  
 RESULT 12  
 ABB80532  
 ID ABB80532 standard; peptide; 11 AA.  
 AC ABB80532;  
 XX  
 DT 08-OCT-2002 (first entry)  
 XX  
 DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #12.  
 XX  
 KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
 KW virucide.  
 XX  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers



```

XX  Sequence  11 AA;
Query Match      82.1%; Score 46; DB 23; Length 11;
Best Local Similarity 90.9%; Pred. No. 0.017;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 EEVVPXGMHYS 11
        ||||| 111
DB      1 EEVVPXGGHYS 11

RESULT 13
ABB80521
ID ABB80521 standard; peptide; 11 AA.
XX
XX ABB80521;
XX
XX
XX 08-OCT-2002 (first entry)
XX
XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #1.
XX
XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
XX virucide.
XX
XX Synthetic.
XX
XX
XX OS
XX
XX Location/Qualifiers
FH
FH Key 1
FT Modified-site 1 /note= "N-terminal acetyl"
FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
FT residue 7"
FT Modified-site 11 /note= "C-terminal amide"
FT
FT
XX

```

RESULT	14	
ABB80522		
ID	ABB80522 standard; peptide; 11 AA.	
XX		
AC	ABB80522;	
XX		
DT	08-OCT-2002 (first entry)	
XX		
DE	Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #2.	
XX		
KW	Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;	
KW	virucide.	
XX		
OS	Synthetic.	
XX		
PH	Key	Location/Qualifiers
FT	Modified-site	1
FT		/note= "N-terminal acetyl"
FT	Modified-site	6
FT		/note= "Norvalyl carbonyl forming keto-amide linkage with residue 7"
FT		
FT	Misc-difference	9
FT		/note= "D-form residue"
FT		
FT	Modified-site	11
FT		/note= "C-terminal amide"
XX		
PN	WO200208251-A2.	
XX		
PD	31-JAN-2002.	
XX		
PF	19-JUL-2001; 2001WO-US23169.	
XX		
PR	21-JUL-2000; 2000US-220101P.	
XX		

PA (CORV-) CORVAS INT INC.  
 XX Lim-wilby M, Levy OE, Brunck TK;  
 PI WPI; 2002-361643/39.  
 XX  
 XX Novel peptide compound having hepatitis C virus protease inhibitory  
 PT activity useful for treating disorders associated with hepatitis C  
 PT virus protease  
 XX  
 XX Claim 17; Page 64; 69pp; English.  
 PS  
 XX The sequence represents a peptide compound of the invention having  
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
 CC invention are alpha-ketoamide peptide analogues. The peptides have  
 CC virucide activity, and are useful for treating and in the manufacture of  
 CC a medicament to treat disorders associated with HCV protease. A  
 CC pharmaceutical composition comprising the peptide as an active ingredient  
 CC is useful for treating disorders associated with hepatitis C virus.  
 XX  
 XX Sequence 11 AA;  
 SQ  
 Query Match 80.4%; Score 45; DB 23; Length 11;  
 Best Local Similarity 90.9%; Pred. No. 0.026;  
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 EEVVPXGMHYS 11  
 Db 1 EEVVPXGMSYS 11  
 RESULT 15  
 ABB80524  
 ID ABB80524 standard; peptide; 11 AA.  
 XX  
 AC ABB80524;  
 XX  
 DT 08-OCT-2002 (first entry)  
 XX  
 DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #4.  
 XX  
 KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
 KW virucide.  
 XX  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Modified-site 1 /note= "N-terminal acetyl"  
 FT Modified-site 6  
 FT Misc-difference 9 /note= "Norvalyl carbonyl forming keto-amide linkage with  
 FT residue 7"  
 FT Modified-site 11 /note= "D-form residue"  
 FT Modified-site 11 /note= "C-terminal amide"  
 XX  
 PN WO200208251-A2.  
 XX  
 PD 31-JAN-2002.  
 XX  
 PF 19-JUL-2001; 2001WO-US23169.  
 XX  
 PR 21-JUL-2000; 2000US-220101P.  
 XX  
 PA (CORV-) CORVAS INT INC.  
 XX  
 PI Lim-wilby M, Levy OE, Brunck TK;  
 XX  
 DR WPI; 2002-361643/39.  
 XX  
 XX Novel peptide compound having hepatitis C virus protease inhibitory  
 PT activity useful for treating disorders associated with hepatitis C

PT virus protease  
 XX Claim 17; Page 64; 69pp; English.  
 PS  
 XX The sequence represents a peptide compound of the invention having  
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
 CC invention are alpha-ketoamide peptide analogues. The peptides have  
 CC virucide activity, and are useful for treating and in the manufacture of  
 CC a medicament to treat disorders associated with HCV protease. A  
 CC pharmaceutical composition comprising the peptide as an active ingredient  
 CC is useful for treating disorders associated with hepatitis C virus.  
 XX  
 XX Sequence 11 AA;  
 SQ  
 Query Match 80.4%; Score 45; DB 23; Length 11;  
 Best Local Similarity 90.9%; Pred. No. 0.026;  
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 EEVVPXGMHYS 11  
 Db 1 EEVVPXGMDYS 11  
 RESULT 16  
 ABB80525  
 ID ABB80525 standard; peptide; 11 AA.  
 XX  
 AC ABB80525;  
 XX  
 DT 08-OCT-2002 (first entry)  
 XX  
 DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #5.  
 XX  
 KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
 KW virucide.  
 XX  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Modified-site 1 /note= "N-terminal acetyl"  
 FT Modified-site 6  
 FT Misc-difference 8 /note= "Norvalyl carbonyl forming keto-amide linkage with  
 FT residue 7"  
 FT Modified-site 11 /note= "D-form residue"  
 FT Modified-site 11 /note= "C-terminal amide"  
 XX  
 PN WO200208251-A2.  
 XX  
 PD 31-JAN-2002.  
 XX  
 PF 19-JUL-2001; 2001WO-US23169.  
 XX  
 PR 21-JUL-2000; 2000US-220101P.  
 XX  
 PA (CORV-) CORVAS INT INC.  
 XX  
 PI Lim-wilby M, Levy OE, Brunck TK;  
 XX  
 DR WPI; 2002-361643/39.  
 XX  
 XX Novel peptide compound having hepatitis C virus protease inhibitory  
 PT activity useful for treating disorders associated with hepatitis C  
 PT virus protease  
 XX  
 XX Claim 17; Page 64; 69pp; English.  
 PS  
 XX The sequence represents a peptide compound of the invention having  
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
 CC invention are alpha-ketoamide peptide analogues. The peptides have  
 CC virucide activity, and are useful for treating and in the manufacture of

CC a medicament to treat disorders associated with HCV protease. A  
CC pharmaceutical composition comprising the peptide as an active ingredient  
CC is useful for treating disorders associated with hepatitis C virus.

XX Sequence 11 AA;

Query Match 80.4%; Score 45; DB 23; Length 11;

Best Local Similarity 90.9%; Pred. No. 0.026;

Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXGMHYS 11

DB 1 EEVVPXGMSYS 11

RESULT 17

ABB80526  
ID ABB80526 standard; peptide; 11 AA.

XX ABB80526;

DT 08-OCT-2002 (first entry)

DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #6.

XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;

KW virucide.

XX Synthetic.

FT Key Location/Qualifiers

FT Modified-site 1

FT /note= "N-terminal acetyl"

FT Modified-site 6

FT /note= "Norvalyl carbonyl forming keto-amide linkage with residue 7"

FT Misc-difference 8

FT /note= "D-form residue"

FT Misc-difference 9

FT /note= "D-form residue"

FT Modified-site 11

FT /note= "C-terminal amide"

XX WO200208251-A2.

PN 31-JAN-2002.

XX 19-JUL-2001; 2001WO-US23169.

XX 21-JUL-2000; 2000US-220101P.

XX (CORV-) CORVAS INT INC.

XX Lim-wilby M, Levy OE, Brunck TK;

XX WPI; 2002-361643/39.

XX Novel peptide compound having hepatitis C virus protease inhibitory  
PT activity useful for treating disorders associated with hepatitis C  
PT virus protease

XX Claim 17; Page 64; 69pp; English.

XX The sequence represents a peptide compound of the invention having  
CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
CC invention are alpha-ketoamide peptide analogues. The peptides have  
CC virucide activity, and are useful for treating and in the manufacture of  
CC a medicament to treat disorders associated with HCV protease. A  
CC pharmaceutical composition comprising the peptide as an active ingredient  
CC is useful for treating disorders associated with hepatitis C virus.

XX Sequence 11 AA;

Query Match 80.4%; Score 45; DB 23; Length 11;

Best Local Similarity 90.9%; Pred. No. 0.026;

Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXGMHYS 11

DB 1 EEVVPXGMSYS 11

RESULT 18

ABB80528

ID ABB80528 standard; peptide; 11 AA.

XX ABB80528;

XX 08-OCT-2002 (first entry)

DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #8.

XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;

KW virucide.

XX Synthetic.

FT Key Location/Qualifiers

FT Modified-site 1

FT /note= "N-terminal acetyl"

FT Modified-site 6

FT /note= "Norvalyl carbonyl forming keto-amide linkage with residue 7"

FT Misc-difference 8

FT /note= "D-form residue"

FT Modified-site 11

FT /note= "C-terminal amide"

XX WO200208251-A2.

PN 31-JAN-2002.

XX 19-JUL-2001; 2001WO-US23169.

XX 21-JUL-2000; 2000US-220101P.

XX (CORV-) CORVAS INT INC.

XX Lim-wilby M, Levy OE, Brunck TK;

XX WPI; 2002-361643/39.

XX Novel peptide compound having hepatitis C virus protease inhibitory  
PT activity useful for treating disorders associated with hepatitis C  
PT virus protease

XX Claim 17; Page 64; 69pp; English.

XX The sequence represents a peptide compound of the invention having  
CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
CC invention are alpha-ketoamide peptide analogues. The peptides have  
CC virucide activity, and are useful for treating and in the manufacture of  
CC a medicament to treat disorders associated with HCV protease. A  
CC pharmaceutical composition comprising the peptide as an active ingredient  
CC is useful for treating disorders associated with hepatitis C virus.

XX Sequence 11 AA;

Query Match 80.4%; Score 45; DB 23; Length 11;

Best Local Similarity 90.9%; Pred. No. 0.026;

Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXGMHYS 11

DB 1 EEVVPXGMDYS 11

RESULT 19  
 ABB80529  
 ID ABB80529 standard; peptide; 11 AA.  
 AC ABB80529;  
 XX  
 DT 08-OCT-2002 (first entry)  
 XX  
 DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #9.  
 XX  
 KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
 KW virucide.  
 XX  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Modified-site 1  
 FT Modified-site 6 /note= "N-terminal acetyl"  
 FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with  
 FT residue 7"  
 FT Misc-difference 8  
 FT /note= "D-form residue"  
 FT Modified-site 8  
 FT /note= "Oxymethionine"  
 FT Modified-site 11  
 FT /note= "C-terminal amide"  
 XX  
 XX WO200208251-A2.  
 XX PN  
 XX PD 31-JAN-2002.  
 XX PF 19-JUL-2001; 2001WO-US23169.  
 XX PR 21-JUL-2000; 2000US-220101P.  
 XX PA (CORV-) CORVAS INT INC.  
 XX PI Lim-wilby M, Levy OE, Brunck TK;  
 XX WPI; 2002-361643/39.  
 XX  
 PT Novel peptide compound having hepatitis C virus protease inhibitory  
 PT activity useful for treating disorders associated with hepatitis C  
 PT virus protease  
 XX  
 PS Claim 17; Page 64; 69pp; English.  
 XX  
 CC The sequence represents a peptide compound of the invention having  
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
 CC invention are alpha-ketoamide peptide analogues. The peptides have  
 CC virucide activity, and are useful for treating and in the manufacture of  
 CC a medicament to treat disorders associated with HCV protease. A  
 CC pharmaceutical composition comprising the peptide as an active ingredient  
 CC is useful for treating disorders associated with hepatitis C virus.  
 XX  
 SQ Sequence 11 AA;  
 Query Match 80.4%; Score 45; DB 23; Length 11;  
 Best Local Similarity 90.9%; Pred. No. 0.026;  
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 EEVVPXGMHYS 11  
 Db | | | | | | | | | |  
 1 EEVVPXGMHYS 11  
 RESULT 20  
 ABB80559  
 ID ABB80559 standard; peptide; 11 AA.  
 AC ABB80559;  
 XX  
 DT 08-OCT-2002 (first entry)  
 XX  
 DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #41.  
 XX  
 KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
 KW virucide.

XX 08-OCT-2002 (first entry)  
 XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #39.  
 DE Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
 KW virucide.  
 KW virucide.  
 XX  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Modified-site 1  
 FT Modified-site 6 /note= "N-terminal acetyl"  
 FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with  
 FT residue 7"  
 FT Misc-difference 8  
 FT /note= "D-form residue"  
 FT Modified-site 8  
 FT /note= "Oxymethionine"  
 FT Modified-site 11  
 FT /note= "C-terminal amide"  
 XX  
 XX WO200208251-A2.  
 XX PN  
 XX PD 31-JAN-2002.  
 XX PF 19-JUL-2001; 2001WO-US23169.  
 XX PR 21-JUL-2000; 2000US-220101P.  
 XX PA (CORV-) CORVAS INT INC.  
 XX PI Lim-wilby M, Levy OE, Brunck TK;  
 XX WPI; 2002-361643/39.  
 XX  
 PT Novel peptide compound having hepatitis C virus protease inhibitory  
 PT activity useful for treating disorders associated with hepatitis C  
 PT virus protease  
 XX  
 PS Claim 17; Page 65; 69pp; English.  
 XX  
 CC The sequence represents a peptide compound of the invention having  
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
 CC invention are alpha-ketoamide peptide analogues. The peptides have  
 CC virucide activity, and are useful for treating and in the manufacture of  
 CC a medicament to treat disorders associated with HCV protease. A  
 CC pharmaceutical composition comprising the peptide as an active ingredient  
 CC is useful for treating disorders associated with hepatitis C virus.  
 XX  
 SQ Sequence 11 AA;  
 Query Match 80.4%; Score 45; DB 23; Length 11;  
 Best Local Similarity 90.9%; Pred. No. 0.026;  
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 EEVVPXGMHYS 11  
 Db | | | | | | | | | |  
 1 EEVVPXGMHYS 11  
 RESULT 21  
 ABB80561  
 ID ABB80561 standard; peptide; 11 AA.  
 XX  
 AC ABB80561;  
 XX  
 DT 08-OCT-2002 (first entry)  
 XX  
 DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #41.  
 XX  
 KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;

[illegible]

FT Modified-site 11 /note= "C-terminal amide"  
 FT  
 PN WO200208251-A2.  
 XX  
 PD 31-JAN-2002.  
 XX  
 PF 19-JUL-2001; 2001WO-US23169.  
 XX  
 PR 21-JUL-2000; 2000US-220101P.  
 XX  
 PA (CORV-) CORVAS INT INC.  
 PI Lim-wilby M, Levy OE, Brunck TK;  
 XX  
 PI WPI; 2002-361643/39.  
 XX  
 DR Novel peptide compound having hepatitis C virus protease inhibitory  
 PT activity useful for treating disorders associated with hepatitis C  
 PT virus protease  
 XX  
 PS Claim 17; Page 65; 69pp; English.  
 XX  
 CC The sequence represents a peptide compound of the invention having  
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
 CC invention are alpha-ketoamide peptide analogues. The peptides have  
 CC virucide activity, and are useful for treating and in the manufacture of  
 CC a medicament to treat disorders associated with HCV protease. A  
 CC pharmaceutical composition comprising the peptide as an active ingredient  
 CC is useful for treating disorders associated with hepatitis C virus.  
 XX  
 SQ Sequence 11 AA;

Query Match 80.4%; Score 45; DB 23; Length 11;  
 Best Local Similarity 90.9%; Pred. No. 0.026;  
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 EEVVPXGMHYS 11  
 Db 1 EEVVPXGMHYS 11  
 |||||

RESULT 24  
 ABB80564  
 ID ABB80564 standard; peptide; 11 AA.  
 AC  
 AC ABB80564;  
 XX  
 XX 08-OCT-2002 (first entry)  
 XX  
 DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #44.  
 XX  
 KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
 KW virucide.  
 XX  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Modified-site 1 /note= "N-terminal acetyl"  
 FT Modified-site 6 /note= "Leucyl carbonyl forming keto-amide linkage with  
 FT Modified-site 11 residue 7"  
 FT Modified-site 11 /note= "C-terminal amide"  
 XX  
 PN WO200208251-A2.  
 XX  
 PD 31-JAN-2002.  
 XX  
 PF 19-JUL-2001; 2001WO-US23169.  
 XX  
 PR 21-JUL-2000; 2000US-220101P.

XX  
 PA (CORV-) CORVAS INT INC.  
 XX  
 PI Lim-wilby M, Levy OE, Brunck TK;  
 XX  
 DR WPI; 2002-361643/39.  
 XX  
 PT Novel peptide compound having hepatitis C virus protease inhibitory  
 PT activity useful for treating disorders associated with hepatitis C  
 PT virus protease  
 XX  
 PS Claim 17; Page 65; 69pp; English.  
 XX  
 CC The sequence represents a peptide compound of the invention having  
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
 CC invention are alpha-ketoamide peptide analogues. The peptides have  
 CC virucide activity, and are useful for treating and in the manufacture of  
 CC a medicament to treat disorders associated with HCV protease. A  
 CC pharmaceutical composition comprising the peptide as an active ingredient  
 CC is useful for treating disorders associated with hepatitis C virus.  
 XX  
 SQ Sequence 11 AA;

Query Match 80.4%; Score 45; DB 23; Length 11;  
 Best Local Similarity 90.9%; Pred. No. 0.026;  
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 EEVVPXGMHYS 11  
 Db 1 EEVVPXGMHYS 11  
 |||||

RESULT 25  
 ABB80565  
 ID ABB80565 standard; peptide; 11 AA.  
 XX  
 AC ABB80565;  
 XX  
 XX 08-OCT-2002 (first entry)  
 XX  
 DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #45.  
 XX  
 KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
 KW virucide.  
 XX  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Modified-site 1 /note= "N-terminal acetyl"  
 FT Modified-site 6 /note= "Norleucyl carbonyl forming keto-amide linkage  
 FT Modified-site 11 with residue 7"  
 FT Modified-site 11 /note= "C-terminal amide"  
 XX  
 PN WO200208251-A2.  
 XX  
 PD 31-JAN-2002.  
 XX  
 PF 19-JUL-2001; 2001WO-US23169.  
 XX  
 PR 21-JUL-2000; 2000US-220101P.  
 XX  
 PA (CORV-) CORVAS INT INC.  
 XX  
 PI Lim-wilby M, Levy OE, Brunck TK;  
 XX  
 DR WPI; 2002-361643/39.  
 XX  
 PT Novel peptide compound having hepatitis C virus protease inhibitory  
 PT activity useful for treating disorders associated with hepatitis C  
 PT virus protease

Wed Jun 11 15:42:08 2003

XX Claim 17; Page 65; 69pp; English.  
 XX The sequence represents a peptide compound of the invention having  
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
 CC invention are alpha-ketoamide peptide analogues. The peptides have  
 CC virucide activity, and are useful for treating and in the manufacture of  
 CC a medicament to treat disorders associated with HCV protease. A  
 CC pharmaceutical composition comprising the peptide as an active ingredient  
 CC is useful for treating disorders associated with hepatitis C virus.

XX SQ Sequence 11 AA;  
 Query Match 80.4%; Score 45; DB 23; Length 11;  
 Best Local Similarity 90.9%; Pred. No. 0.026;  
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 EEVVPXGMHYS 11  
 |||||  
 Db 1 EEVVPXGMSYS 11

Search completed: June 10, 2003, 13:39:08  
 Job time : 31.3571 secs





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2003, 13:31:45 ; Search time 9.64286 Seconds  
(without alignments)  
33.564 Million cell updates/sec

Title: US-09-909-164-11  
Perfect score: 56  
Sequence: 1 EEVVPXGMHYS 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues  
Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 75 summaries

Database : Issued\_Patents\_AA.\*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pap.\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pap.\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pap.\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pap.\*  
5: /cgn2\_6/ptodata/1/iaa/PTUS\_COMB.pap.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pap.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	39	69.6	1037	4	US-09-134-001C-4794
2	34	60.7	600	2	US-08-821-119-19
3	34	60.7	600	2	US-08-821-118-2
4	33	58.9	747	4	US-09-724-864-36
5	32	57.1	70	4	US-09-134-001C-3950
6	32	57.1	102	2	US-08-580-988A-23
7	32	57.1	126	3	US-08-879-995A-3
8	32	57.1	126	3	US-09-215-096-3
9	32	57.1	152	3	US-08-460-694-4
10	32	57.1	152	3	US-08-460-744-4
11	32	57.1	152	3	US-07-667-711B-4
12	32	57.1	173	1	US-08-193-977-7
13	32	57.1	189	2	US-08-464-517-21
14	32	57.1	189	2	US-08-246-361A-21
15	32	57.1	189	3	US-08-463-772-21
16	32	57.1	189	5	PCT-US93-05000-21
17	32	57.1	236	2	US-08-464-517-22
18	32	57.1	236	2	US-08-246-361A-22
19	32	57.1	236	5	PCT-US93-05000-22
20	32	57.1	236	5	PCT-US93-05000-22
21	32	57.1	280	3	US-08-464-517-6
22	32	57.1	280	3	US-08-463-772-6
23	32	57.1	289	2	US-08-246-361A-4
24	32	57.1	289	5	PCT-US93-05000-4
25	32	57.1	291	5	PCT-US93-05000-6
26	32	57.1	292	2	US-08-464-517-23
27	32	57.1	292	2	US-08-246-361A-6

28	57.1	292	2	US-08-246-361A-23	Sequence 23, Appl
29	57.1	292	3	US-08-463-772-23	Sequence 23, Appl
30	57.1	292	5	PCT-US93-05000-23	Sequence 23, Appl
31	57.1	295	1	US-07-947-120-8	Sequence 8, Appl
32	57.1	295	1	US-08-472-893A-8	Sequence 8, Appl
33	57.1	295	2	US-08-460-694-2	Sequence 2, Appl
34	57.1	295	2	US-08-464-517-19	Sequence 19, Appl
35	57.1	295	2	US-08-464-517-20	Sequence 20, Appl
36	57.1	295	2	US-08-246-361A-19	Sequence 19, Appl
37	57.1	295	2	US-08-246-361A-20	Sequence 20, Appl
38	57.1	295	3	US-08-463-772-19	Sequence 19, Appl
39	57.1	295	3	US-08-463-772-20	Sequence 20, Appl
40	57.1	295	3	US-08-460-744-2	Sequence 2, Appl
41	57.1	295	3	US-07-667-711B-2	Sequence 2, Appl
42	57.1	295	3	US-08-947-432-8	Sequence 8, Appl
43	57.1	295	5	PCT-US93-05000-2	Sequence 2, Appl
44	57.1	295	5	PCT-US93-05000-19	Sequence 19, Appl
45	57.1	295	5	PCT-US93-05000-20	Sequence 20, Appl
46	57.1	309	2	US-08-464-517-4	Sequence 4, Appl
47	57.1	309	3	US-08-463-772-4	Sequence 4, Appl
48	57.1	615	2	US-08-663-566A-17	Sequence 17, Appl
49	57.1	615	2	US-08-023-610-17	Sequence 17, Appl
50	57.1	615	2	US-08-288-065A-17	Sequence 17, Appl
51	57.1	615	2	US-08-362-240A-17	Sequence 17, Appl
52	57.1	615	5	PCT-US95-10245-17	Sequence 17, Appl
53	57.1	618	2	US-08-770-761A-3	Sequence 3, Appl
54	57.1	647	2	US-08-770-761A-8	Sequence 8, Appl
55	57.1	660	2	US-08-770-761A-2	Sequence 2, Appl
56	57.1	662	2	US-08-770-761A-5	Sequence 5, Appl
57	57.1	705	2	US-08-770-761A-7	Sequence 7, Appl
58	57.1	819	2	US-08-464-517-7	Sequence 7, Appl
59	57.1	819	2	US-08-246-361A-7	Sequence 7, Appl
60	57.1	819	3	US-08-463-772-7	Sequence 7, Appl
61	57.1	819	5	PCT-US93-05000-7	Sequence 7, Appl
62	56.2	501	2	US-08-408-095-31	Sequence 31, Appl
63	55.4	59	4	US-08-963-851-14	Sequence 14, Appl
64	55.4	65	6	5177197-51	Patent No. 5177197
65	55.4	410	6	5177197-1	Patent No. 5177197
66	55.4	447	4	US-08-961-083-182	Sequence 182, Appl
67	55.4	502	4	US-09-342-647-4	Sequence 4, Appl
68	55.4	529	4	US-09-240-639-4	Sequence 4, Appl
69	55.4	622	2	US-08-459-146-2	Sequence 2, Appl
70	55.4	622	2	US-08-459-065-2	Sequence 2, Appl
71	55.4	630	4	US-09-342-647-2	Sequence 2, Appl
72	55.4	667	4	US-09-342-647-28	Sequence 28, Appl
73	55.4	738	1	US-08-530-010-3	Sequence 3, Appl
74	55.4	738	1	US-08-530-010-5	Sequence 5, Appl
75	55.4	738	1	US-08-530-010-7	Sequence 7, Appl

ALIGNMENTS

RESULT 1  
US-09-134-001C-4794  
; Sequence 4794, Application US/09134001C  
; Patent No. 6380370  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: GTC-007  
; CURRENT APPLICATION NUMBER: US/09/134,001C  
; CURRENT FILING DATE: 1998-08-13  
; PRIOR APPLICATION NUMBER: US 60/064,964  
; PRIOR FILING DATE: 1997-11-08  
; PRIOR APPLICATION NUMBER: US 60/055,779  
; PRIOR FILING DATE: 1997-08-14  
; NUMBER OF SEQ ID NOS: 5674  
; SEQ ID NO 4794  
; LENGTH: 1037  
; TYPE: PRT  
; ORGANISM: Staphylococcus epidermidis

US-09-134-001C-4794

Query Match 69.6%; Score 39; DB 4; Length 1037;  
Best Local Similarity 63.6%; Pred. No. 16;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMHYS 11  
:|||||:|||||  
DB 199 KEVSNGLHYS 209

## RESULT 2

US-08-821-119-19  
; Sequence 19, Application US/08821119  
; Patent No. 5821104  
; GENERAL INFORMATION:  
; APPLICANT: Holm, Kaj Andre  
; APPLICANT: Rasmussen, Grethe  
; APPLICANT: Haikier, Torben  
; APPLICANT: Lehmbeck, Jan  
; TITLE OF INVENTION: Tripeptidyl Amino peptidase  
; NUMBER OF SEQUENCES: 23  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: No. 5821104o No. 5821104disk of No. 5821104th America, Inc.  
; STREET: 405 Lexington Avenue  
; CITY: New York  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10174  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/821,119  
; FILING DATE: 19-MAR-1997  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lambiris, Elias J  
; REGISTRATION NUMBER: 33,728  
; REFERENCE/DOCKET NUMBER: 4107.204-US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-867-0123  
; TELEFAX: 212-878-9655  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 19:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 600 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; FRAGMENT TYPE: internal  
US-08-821-119-19

Query Match 60.7%; Score 34; DB 2; Length 600;  
Best Local Similarity 75.0%; Pred. No. 82;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 VPXGMHYS 11  
:|||||:|||||  
DB 31 VPKGWHYS 38

## RESULT 3

US-08-821-118-2  
; Sequence 2, Application US/08821118  
; Patent No. 5989889  
; GENERAL INFORMATION:  
; APPLICANT: Rev, Michael  
; APPLICANT: Golightly, Elizabeth  
; TITLE OF INVENTION: NUCLEIC ACIDS ENCODING POLYPEPTIDES

; TITLE OF INVENTION: HAVING TRIPEPTIDE AMINOPEPTIDASE  
; TITLE OF INVENTION: ACTIVITY  
; NUMBER OF SEQUENCES: 7  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: No. 5989889o No. 5989889disk of No. 5989889th America, Inc.  
; STREET: 405 Lexington Avenue  
; CITY: New York  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10174  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/821,118  
; FILING DATE: 19-MAR-1997  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lambiris, Elias J  
; REGISTRATION NUMBER: 33,728  
; REFERENCE/DOCKET NUMBER: 4107.400-US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-867-0123  
; TELEFAX: 212-878-9655  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 600 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; FRAGMENT TYPE: internal  
US-08-821-118-2

Query Match 60.7%; Score 34; DB 2; Length 600;  
Best Local Similarity 75.0%; Pred. No. 82;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 VPXGMHYS 11  
:|||||:|||||  
DB 31 VPKGWHYS 38

## RESULT 4

US-09-724-864-36  
; Sequence 36, Application US/09724864  
; Patent No. 6380362  
; GENERAL INFORMATION:  
; APPLICANT: Watson, James D  
; APPLICANT: Murison, James G  
; TITLE OF INVENTION: Polynucleotides, polypeptides expressed  
; FILE REFERENCE: by the polynucleotides and methods for their use.  
; FILE REFERENCE: 11000.1050U1  
; CURRENT APPLICATION NUMBER: US/09/724,864  
; CURRENT FILING DATE: 2000-11-28  
; PRIOR APPLICATION NUMBER: U.S. No. 6380362 60/171,678  
; PRIOR FILING DATE: 1999-12-23  
; NUMBER OF SEQ ID NOS: 72  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 36  
; LENGTH: 747  
; TYPE: PRT  
; ORGANISM: Rat  
US-09-724-864-36

Query Match 58.9%; Score 33; DB 4; Length 747;  
Best Local Similarity 71.4%; Pred. No. 1.6e+02;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 PXGMHYS 11

Db 627 PGLHYS 633  
| | | | |

## RESULT 5

US-09-134-001C-3950  
; Sequence 3950, Application US/09134001C  
; Patent No. 6380370  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
; FILE REFERENCE: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS  
; CURRENT APPLICATION NUMBER: US/09/134,001C  
; CURRENT FILING DATE: 1998-08-13  
; PRIOR APPLICATION NUMBER: US 60/064,964  
; PRIOR FILING DATE: 1997-11-08  
; PRIOR APPLICATION NUMBER: US 60/055,779  
; PRIOR FILING DATE: 1997-08-14  
; NUMBER OF SEQ ID NOS: 5674  
; SEQ ID NO 3950  
; LENGTH: 70  
; TYPE: PRT  
; ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-3950

Query Match 57.1%; Score 32; DB 4; Length 70;

Best Local Similarity 62.3%; Pred. No. 19;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 VPXGMHYS 11

Db 36 MPKGFHYS 43

## RESULT 6

US-08-580-988A-23  
; Sequence 23, Application US/08580988A  
; Patent No. 5856161  
; GENERAL INFORMATION:  
; APPLICANT: Aggarwal et al.  
; TITLE OF INVENTION: Tumor Necrosis Factor  
; TITLE OF INVENTION: Receptor-I-Associated Protein Kinase And Methods  
; TITLE OF INVENTION: For Its Use  
; NUMBER OF SEQUENCES: 27  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dr. Benjamin A. Adler  
; STREET: 8011 Candle Lane  
; CITY: Houston  
; STATE: Texas  
; COUNTRY: USA  
; ZIP: 77071  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 1.44 Mb floppy disk  
; COMPUTER: Apple Macintosh  
; OPERATING SYSTEM: Macintosh  
; SOFTWARE: Microsoft Word for Macintosh  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/580,988A  
; FILING DATE: January 3, 1996  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Benjamin Aaron Adler, Ph.D., J.D.  
; REGISTRATION NUMBER: 35,423  
; REFERENCE/DOCKET NUMBER: D5721CIP2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 713-777-2321  
; TELEFAX: 713-777-6908  
; INFORMATION FOR SEQ ID NO: 23:  
; SEQUENCE CHARACTERISTICS:

; LENGTH: 102 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE:  
; DESCRIPTION: protein  
; HYPOTHETICAL: no  
; ANTI-SENSE: no  
; FRAGMENT TYPE: internal  
; ORIGINAL SOURCE:  
US-08-580-988A-23

Query Match 57.1%; Score 32; DB 2; Length 102;  
Best Local Similarity 60.0%; Pred. No. 29;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVFXGMHY 10

Db 24 EEVFFLAMNY 33

## RESULT 7

US-08-879-995A-3  
; Sequence 3, Application US/08879995A  
; Patent No. 5985606  
; GENERAL INFORMATION:  
; APPLICANT: Hillman, Jennifer L.  
; APPLICANT: Lal, Preeti  
; APPLICANT: Kaser, Matthew R.  
; TITLE OF INVENTION: HUMAN PREPROTACHYKININ B  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/879,995A  
; FILING DATE: Herewith  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J.  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0326 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-855-0555  
; TELEFAX: 415-845-4166  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 126 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: Genbank  
; CLONE: 163590  
US-08-879-995A-3

Query Match 57.1%; Score 32; DB 2; Length 126;  
Best Local Similarity 66.7%; Pred. No. 36;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMH 9  
 Db 28 EQVPPGGGH 36

## RESULT 8

US-09-215-096-3  
 ; Sequence 3, Application US/09215096  
 ; Patent No. 6008194  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Hillman, Jeunifer L.  
 ; APPLICANT: Lal, Preeti  
 ; APPLICANT: Kaser, Matthew R.  
 ; TITLE OF INVENTION: HUMAN PREPROTACHYKININ B  
 ; NUMBER OF SEQUENCES: 4  
 ; CORRESPONDENCE ADDRESSES:  
 ; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
 ; STREET: 3174 Porter Drive  
 ; CITY: Palo Alto  
 ; STATE: CA  
 ; COUNTRY: USA  
 ; ZIP: 94304  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette  
 ; COMPUTER: IBM Compatible  
 ; OPERATING SYSTEM: DOS  
 ; SOFTWARE: FastSeq for Windows Version 2.0  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/215,096  
 ; FILING DATE:  
 ; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/879,995  
 ; FILING DATE:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Billings, Lucy J.  
 ; REGISTRATION NUMBER: 36,749  
 ; REFERENCE/DOCKET NUMBER: PF-0326 US  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 415-855-0555  
 ; TELEFAX: 415-845-4166  
 ; TELEX:  
 ; INFORMATION FOR SEQ ID NO: 3:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 126 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; IMMEDIATE SOURCE:  
 ; LIBRARY: GenBank  
 ; CLONE: 163590  
 ;

Query Match 57.1%; Score 32; DB 3; Length 126;  
 Best Local Similarity 66.7%; Pred. No. 36;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMH 9  
 Db 28 EQVPPGGGH 36

## RESULT 9

US-08-460-694-4  
 ; Sequence 4, Application US/08460694  
 ; Patent No. 5858655  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Arnold, Andrew  
 ; TITLE OF INVENTION: PRAD1 Cyclin and its cdNA  
 ; NUMBER OF SEQUENCES: 8  
 ; CORRESPONDENCE ADDRESSES:  
 ; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
 ; STREET: 1100 New York Avenue, N.W., Suite 600  
 ;

; CITY: Washington  
 ; STATE: DC  
 ; COUNTRY: USA  
 ; ZIP: 20005  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/460,694  
 ; FILING DATE: 02-JUN-1995  
 ; CLASSIFICATION: 435  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: McConathy, Evelyn H.  
 ; REGISTRATION NUMBER: 35,279  
 ; REFERENCE/DOCKET NUMBER: 0609.4070002  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 202-371-2600  
 ; TELEFAX: 202-371-2540  
 ; INFORMATION FOR SEQ ID NO: 4:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 152 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: not relevant  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: peptide  
 ;

Query Match 57.1%; Score 32; DB 2; Length 152;  
 Best Local Similarity 60.0%; Pred. No. 44;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMHY 10  
 Db 20 EEVFPAMNY 29

RESULT 10  
 US-08-460-744-4  
 ; Sequence 4, Application US/08460744  
 ; Patent No. 6107541  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Arnold, Andrew  
 ; TITLE OF INVENTION: PRAD1 Cyclin and its cdNA  
 ; NUMBER OF SEQUENCES: 8  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
 ; STREET: 1100 New York Avenue, N.W., Suite 600  
 ; CITY: Washington  
 ; STATE: DC  
 ; COUNTRY: USA  
 ; ZIP: 20005  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/460,744  
 ; FILING DATE: 02-JUN-1995  
 ; CLASSIFICATION: 435  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: McConathy, Evelyn H.  
 ; REGISTRATION NUMBER: 35,279  
 ; REFERENCE/DOCKET NUMBER: 0609.4070005  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 202-371-2600  
 ; TELEFAX: 202-371-2540  
 ; INFORMATION FOR SEQ ID NO: 4:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 152 amino acids  
 ; TYPE: amino acid  
 ;

STRANDEDNESS: not relevant  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-460-744-4

Query Match 57.1%; Score 32; DB 3; Length 152;  
Best Local Similarity 60.0%; Pred. No. 44;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMHY 10  
||| | | |  
DB 20 EEVFPPLAMNY 29

RESULT 11  
US-07-667-711B-4  
; Sequence 4, Application US/07667711B  
; Patent No. 6110700  
; GENERAL INFORMATION:  
; APPLICANT: ARNOLD, ANDREW  
; TITLE OF INVENTION: Pradl Cyclin and its cDNA  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
; STREET: 1100 NEW YORK AVE., NW, SUITE 600  
; CITY: WASHINGTON  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/667,711B  
; FILING DATE: 11-MAR-1991  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: MCPHALL, DONALD R.  
; REGISTRATION NUMBER: 35,811  
; REFERENCE/DOCKET NUMBER: 0609.4070000  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 371-2600  
; TELEFAX: (202) 371-2540  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 152 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: not relevant  
; MOLECULE TYPE: peptide  
US-07-667-711B-4

Query Match 57.1%; Score 32; DB 3; Length 152;  
Best Local Similarity 60.0%; Pred. No. 44;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMHY 10  
||| | | |  
DB 20 EEVFPPLAMNY 29

RESULT 12  
US-08-193-977-7  
; Sequence 7, Application US/08193977  
; Patent No. 5625031  
; GENERAL INFORMATION:  
; APPLICANT: WEBSTER, KEVIN R.  
; APPLICANT: COLEMAN, KEVIN G.  
; TITLE OF INVENTION: PEPTIDE INHIBITORS OF THE P33CDK2 AND  
; TITLE OF INVENTION: P34CDK2 CELL CYCLE REGULATORY KINASES AND HUMAN  
; TITLE OF INVENTION: PAPILLOMAVIRUS E7 ONCOPROTEIN

NUMBER OF SEQUENCES: 34  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: REED & ROBINS  
STREET: 635 BRYANT STREET  
CITY: PALO ALTO  
STATE: CALIFORNIA  
COUNTRY: UNITED STATES OF AMERICA  
ZIP: 94301  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/193,977  
FILING DATE: 08-FEB-1994  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: ROBINS, ROBERTA L.  
REGISTRATION NUMBER: 33,208  
REFERENCE/DOCKET NUMBER: 5998-0016  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 617-8999  
TELEFAX: (415) 327-3231  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 173 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-193-977-7

Query Match 57.1%; Score 32; DB 1; Length 173;  
Best Local Similarity 60.0%; Pred. No. 51;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMHY 10  
||| | | |  
DB 55 EEVFPPLAMNY 64

RESULT 13  
US-08-464-517-21  
; Sequence 21, Application US/08464517  
; Patent No. 5869640  
; GENERAL INFORMATION:  
; APPLICANT: BEACH, David H.  
; TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO  
; NUMBER OF SEQUENCES: 50  
; CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD  
STREET: 60 State Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII(text)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/464,517  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/963,308  
FILING DATE: 16-OCT-1992  
APPLICATION NUMBER: US 07/888,178  
FILING DATE: 26-MAY-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/701,514

;; FILING DATE: 16-MAY-1991  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Matthew P. Vincent  
;; REGISTRATION NUMBER: 36,709  
;; REFERENCE/DOCKET NUMBER: MII-004C  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (617) 227-7400  
;; TELEFAX: (617) 227-5941  
;; INFORMATION FOR SEQ ID NO: 21:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 189 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
US-08-464-517-21

Query Match 57.1%; Score 32; DB 2; Length 189;  
Best Local Similarity 60.0%; Pred. No. 56;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
QY 1 EEVVPXGMHY 10  
Db 74 EEVFLPMY 83

RESULT 14  
US-08-246-361A-21  
; Sequence 21, Application US/08246361A  
; Patent No. 5998582  
; GENERAL INFORMATION:  
; APPLICANT: BEACH, David H.  
; TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO  
; NUMBER OF SEQUENCES: 50  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD  
; STREET: 60 State Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109

;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: ASCII(text)  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/246,361A  
;; FILING DATE: 19-MAY-1994  
;; CLASSIFICATION: 435  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/963,308  
;; FILING DATE: 16-OCT-1992  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/888,178  
;; FILING DATE: 26-MAY-1992  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/701,514  
;; FILING DATE: 16-MAY-1991  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Matthew P. Vincent  
;; REGISTRATION NUMBER: 36,709  
;; REFERENCE/DOCKET NUMBER: MII-004C  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (617) 227-7400  
;; TELEFAX: (617) 227-5941  
;; INFORMATION FOR SEQ ID NO: 21:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 189 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide

US-08-246-361A-21

Query Match 57.1%; Score 32; DB 2; Length 189;  
Best Local Similarity 60.0%; Pred. No. 56;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMHY 10  
Db 74 EEVFLPMY 83

RESULT 15  
US-08-463-772-21  
; Sequence 21, Application US/08463772  
; Patent No. 6066501  
; GENERAL INFORMATION:  
; APPLICANT: BEACH, David H.  
; TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO  
; NUMBER OF SEQUENCES: 50  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD  
; STREET: 60 State Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109

;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: ASCII(text)  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/463,772  
;; FILING DATE:  
;; CLASSIFICATION: 435  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/963,308  
;; FILING DATE: 16-OCT-1992  
;; APPLICATION NUMBER: US 07/888,178  
;; FILING DATE: 26-MAY-1992  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/701,514  
;; FILING DATE: 16-MAY-1991  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Matthew P. Vincent  
;; REGISTRATION NUMBER: 36,709  
;; REFERENCE/DOCKET NUMBER: MII-004C  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (617) 227-7400  
;; TELEFAX: (617) 227-5941  
;; INFORMATION FOR SEQ ID NO: 21:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 189 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
US-08-463-772-21

Query Match 57.1%; Score 32; DB 3; Length 189;  
Best Local Similarity 60.0%; Pred. No. 56;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMHY 10  
Db 74 EEVFLPMY 83

RESULT 16  
PCT-US93-05000-21  
; Sequence 21, Application PC/TUS9305000  
; GENERAL INFORMATION:  
; APPLICANT: MITOTIX

;; TITLE OF INVENTION: D-Type Cyclin and Uses Related Thereto  
;; NUMBER OF SEQUENCES: 42  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
;; STREET: Two Militia Drive  
;; CITY: Lexington  
;; STATE: Massachusetts  
;; COUNTRY: US  
;; ZIP: 02173  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: Patentin Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: PCT/US93/05000  
;; FILING DATE: 19930525  
;; CLASSIFICATION:  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US/07/888,178  
;; FILING DATE: 26-MAY-1992  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Granahan, Patricia  
;; REGISTRATION NUMBER: 32,227  
;; REFERENCE/DOCKET NUMBER: CSHL91-02A  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 616-861-6240  
;; TELEFAX: 616-861-9540  
;; INFORMATION FOR SEQ ID NO: 21:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 189 amino acids  
;; TYPE: AMINO ACID  
;; TOPOLOGY: unknown  
;; MOLECULE TYPE: protein  
;; PCT-US93-05000-21

Query Match 57.1%; Score 32; DB 5; Length 189;  
Best Local Similarity 60.0%; Pred. No. 56;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMHY 10  
Db 74 EEVFPPLAMNY 83

RESULT 17  
US-08-464-517-22  
Sequence 22, Application US/08464517  
Patent No. 5869640  
GENERAL INFORMATION:  
APPLICANT: BEACH, David H.  
TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO  
NUMBER OF SEQUENCES: 50  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD  
STREET: 60 State Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII(text)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/464,517  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/963,308  
FILING DATE: 16-OCT-1992  
APPLICATION NUMBER: US 07/888,178

;; FILING DATE: 26-MAY-1992  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/701,514  
;; FILING DATE: 16-MAY-1991  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Matthew P. Vincent  
;; REGISTRATION NUMBER: 36,709  
;; REFERENCE/DOCKET NUMBER: MII-004C  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (617) 227-7400  
;; TELEFAX: (617) 227-5941  
;; INFORMATION FOR SEQ ID NO: 22:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 236 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
;; US-08-464-517-22

Query Match 57.1%; Score 32; DB 2; Length 236;  
Best Local Similarity 60.0%; Pred. No. 72;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMHY 10  
Db 20 EEVFPPLAMNY 29

RESULT 18  
US-08-246-361A-22  
Sequence 22, Application US/08246361A  
Patent No. 5998582  
GENERAL INFORMATION:  
APPLICANT: BEACH, David H.  
TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO  
NUMBER OF SEQUENCES: 50  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD  
STREET: 60 State Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII(text)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/246,361A  
FILING DATE: 19-MAY-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/963,308  
FILING DATE: 16-OCT-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/888,178  
FILING DATE: 26-MAY-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/701,514  
FILING DATE: 16-MAY-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Matthew P. Vincent  
REGISTRATION NUMBER: 36,709  
REFERENCE/DOCKET NUMBER: MII-004C  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 227-7400  
TELEFAX: (617) 227-5941  
INFORMATION FOR SEQ ID NO: 22:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 236 amino acids  
TYPE: amino acid

STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 US-08-246-361A-22

Query Match 57.1%; Score 32; DB 2; Length 236;  
 Best Local Similarity 60.0%; Pred. No. 72;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMHY 10  
 ||| | | |  
 Db 20 EEVFPAMNY 29

## RESULT 19

US-08-463-772-22  
 Sequence 22, Application US/08463772  
 Patent No. 6066501

## GENERAL INFORMATION:

APPLICANT: BEACH, David H.  
 TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO  
 NUMBER OF SEQUENCES: 50  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: LAHIVE & COCKFIELD  
 STREET: 60 State Street  
 CITY: Boston  
 STATE: MA  
 COUNTRY: USA  
 ZIP: 02109

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: ASCII(text)

## CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/463,772  
 FILING DATE:

## CLASSIFICATION: 435

## PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/963,308  
 FILING DATE: 16-OCT-1992  
 APPLICATION NUMBER: US 07/888,178  
 FILING DATE: 26-MAY-1992

## PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/701,514  
 FILING DATE: 16-MAY-1991

## ATTORNEY/AGENT INFORMATION:

NAME: Matthew P. Vincent  
 REGISTRATION NUMBER: 36,709  
 REFERENCE/DOCKET NUMBER: MII-004C  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (617) 227-7400  
 TELEFAX: (617) 227-5941

## INFORMATION FOR SEQ ID NO:

SEQUENCE CHARACTERISTICS:  
 LENGTH: 236 amino acids  
 TYPE: amino acid

## STRANDEDNESS: single

## TOPOLOGY: linear

## MOLECULE TYPE: peptide

US-08-463-772-22

## Query Match

Best Local Similarity 57.1%; Score 32; DB 3; Length 236;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMHY 10  
 ||| | | |  
 Db 20 EEVFPAMNY 29

## RESULT 20

PCT-US93-05000-22

Sequence 22, Application PC/TUS9305000  
 GENERAL INFORMATION:

## APPLICANT: MITOTIX

TITLE OF INVENTION: D-Type Cyclin and Uses Related Thereto  
 NUMBER OF SEQUENCES: 42  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
 STREET: Two Militia Drive  
 CITY: Lexington  
 STATE: Massachusetts  
 COUNTRY: US  
 ZIP: 02173

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25

## CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US93/05000

## FILING DATE: 19930525

## CLASSIFICATION:

## PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/07/888,178

## FILING DATE: 26-MAY-1992

## ATTORNEY/AGENT INFORMATION:

NAME: Granahan, Patricia

## REGISTRATION NUMBER: 32,227

REFERENCE/DOCKET NUMBER: CSHL91-02A

## TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-861-6240

TELEFAX: 616-861-9540

INFORMATION FOR SEQ ID NO: 22:

SEQUENCE CHARACTERISTICS:

LENGTH: 236 amino acids

TYPE: AMINO ACID

TOPOLOGY: unknown

MOLECULE TYPE: protein

PCT-US93-05000-22

## Query Match

Best Local Similarity 57.1%; Score 32; DB 5; Length 236;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMHY 10  
 ||| | | |

Db 20 EEVFPAMNY 29

## RESULT 21

US-08-464-517-6

Sequence 6, Application US/08464517

Patent No. 5869640

## GENERAL INFORMATION:

APPLICANT: BEACH, David H.

TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO

NUMBER OF SEQUENCES: 50

CORRESPONDENCE ADDRESS:

ADDRESSEE: LAHIVE & COCKFIELD

STREET: 60 State Street

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02109

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: ASCII(text)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/464,517

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:



APPLICATION NUMBER: US 07/963,308  
FILING DATE: 16-OCT-1992  
APPLICATION NUMBER: US 07/888,178  
FILING DATE: 26-MAY-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/701,514  
FILING DATE: 16-MAY-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Matthew P. Vincent  
REGISTRATION NUMBER: 36,709  
REFERENCE/DOCKET NUMBER: MII-004C  
TELEPHONE: (617) 227-7400  
TELEFAX: (617) 227-5941  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 280 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-464-517-6

Query Match 57.1%; Score 32; DB 2; Length 280;  
Best Local Similarity 60.0%; Pred. No. 87;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMHY 10  
Db 75 EEVFPAMNY 84

RESULT 22  
US-08-463-772-6  
Sequence 6, Application US/08463772  
Patent No. 6066501  
GENERAL INFORMATION:  
APPLICANT: BEACH, David H.  
TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO  
NUMBER OF SEQUENCES: 50  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD  
STREET: 60 State Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII(text)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/463,772  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/963,308  
FILING DATE: 16-OCT-1992  
APPLICATION NUMBER: US 07/888,178  
FILING DATE: 26-MAY-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/701,514  
FILING DATE: 16-MAY-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Matthew P. Vincent  
REGISTRATION NUMBER: 36,709  
REFERENCE/DOCKET NUMBER: MII-004C  
TELEPHONE: (617) 227-7400  
TELEFAX: (617) 227-5941  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 280 amino acids

TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-463-772-6

Query Match 57.1%; Score 32; DB 3; Length 280;  
Best Local Similarity 60.0%; Pred. No. 87;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMHY 10  
Db 75 EEVFPAMNY 84

RESULT 23  
US-08-246-361A-4  
Sequence 4, Application US/08246361A  
Patent No. 5998582  
GENERAL INFORMATION:  
APPLICANT: BEACH, David H.  
TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO  
NUMBER OF SEQUENCES: 50  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD  
STREET: 60 State Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII(text)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/246,361A  
FILING DATE: 19-MAY-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/963,308  
FILING DATE: 16-OCT-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/888,178  
FILING DATE: 26-MAY-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/701,514  
FILING DATE: 16-MAY-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Matthew P. Vincent  
REGISTRATION NUMBER: 36,709  
REFERENCE/DOCKET NUMBER: MII-004C  
TELEPHONE: (617) 227-7400  
TELEFAX: (617) 227-5941  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 289 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-246-361A-4

Query Match 57.1%; Score 32; DB 2; Length 289;  
Best Local Similarity 60.0%; Pred. No. 90;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMHY 10  
Db 74 EEVFPAMNY 83

RESULT 24  
PCT-US93-05000-4

```

; Sequence 4, Application PC/TUS9305000
; GENERAL INFORMATION:
; APPLICANT: MITOTIX
; TITLE OF INVENTION: D-Type Cyclin and Uses Related Thereto
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: US
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/05000
; FILING DATE: 19930525
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/888,178
; FILING DATE: 26-MAY-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: CSHL91-02A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-9540
; TELEFAX: 616-861-9540
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 289 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; PCT-US93-05000-4

```

```

Query Match          57.1%; Score 32; DB 5; Length 289;
Best Local Similarity 60.0%; Pred. No. 90;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

```

```

QY      1 EEVVPXGMHY 10
        ||| | | |
Db      74 EEVFPPLAMNY 83

```

## RESULT 25

```

PCT-US93-05000-6
; Sequence 6, Application PC/TUS9305000
; GENERAL INFORMATION:
; APPLICANT: MITOTIX
; TITLE OF INVENTION: D-Type Cyclin and Uses Related Thereto
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: US
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/05000
; FILING DATE: 19930525
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/888,178

```

```

; FILING DATE: 26-MAY-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: CSHL91-02A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-9540
; TELEFAX: 616-861-9540
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 291 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; PCT-US93-05000-6

```

```

Query Match          57.1%; Score 32; DB 5; Length 291;
Best Local Similarity 60.0%; Pred. No. 91;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

```

```

QY      1 EEVVPXGMHY 10
        ||| | | |
Db      75 EEVFPPLAMNY 84

```

```

Search completed: June 10, 2003, 13:51:33
Job time : 9.64286 secs

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2003, 13:46:50 ; Search time 15 Seconds  
(without alignments)  
75.710 Million cell updates/sec

Title: US-09-909-164-11

Perfect score: 56

Sequence: 1 ERVVPXGMHYS 11

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 392085 seqs, 103240269 residues

Total number of hits satisfying chosen parameters: 392085

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 75 summaries

#### Database :

Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	33	58.9	567	9 US-10-270-333-126	Sequence 126, App
2	33	58.9	622	9 US-09-738-626-4919	Sequence 4919, App
3	33	58.9	747	9 US-09-866-050A-663	Sequence 663, App
4	33	58.9	3472	9 US-10-027-806-4	Sequence 4, Appli
5	33	58.9	3472	9 US-10-034-623-4	Sequence 4, Appli
6	33	58.9	3472	9 US-10-027-801-4	Sequence 4, Appli
7	32	57.1	254	10 US-09-778-927A-53	Sequence 53, Appl
8	32	57.1	289	9 US-10-024-066-2	Sequence 2, Appli
9	32	57.1	289	9 US-10-024-066-4	Sequence 4, Appli
10	32	57.1	289	10 US-09-919-497-54	Sequence 54, Appl
11	32	57.1	653	9 US-09-820-843A-26	Sequence 26, Appl
12	32	57.1	715	9 US-09-252-088-15	Sequence 16, Appl
13	32	57.1	793	9 US-09-252-088-15	Sequence 15, Appl
14	31.5	56.2	662	9 US-10-047-542-80	Sequence 80, Appl
15	31.5	56.2	847	9 US-09-870-759-52	Sequence 52, Appl
16	31	55.4	59	10 US-09-948-080-14	Sequence 14, Appl
17	31	55.4	73	10 US-09-864-761-40832	Sequence 40832, A
18	31	55.4	192	9 US-10-001-857-119	Sequence 119, App
19	31	55.4	192	9 US-09-986-480-171	Sequence 171, App

#### ALIGNMENTS

RESULT 1  
US-10-270-333-126  
; Sequence 126, Application US/10270333  
; Publication No. US20030092124A1  
; GENERAL INFORMATION:  
; APPLICANT: Craychik, Anibal  
; TITLE OF INVENTION: ISOLATED G-PROTEIN COUPLED RECEPTORS,  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING GPCR PROTEINS, AND USES  
; TITLE OF INVENTION: THEROF AS INSECTICIDAL TARGETS  
; FILE REFERENCE: CLO00733CON  
; CURRENT APPLICATION NUMBER: US/10/270,333  
; CURRENT FILING DATE: 2002-10-15

20	31	55.4	260	10	US-09-815-242-13489	Sequence 13489, A
21	31	55.4	260	10	US-09-815-242-13613	Sequence 13613, A
22	31	55.4	299	10	US-09-815-242-10697	Sequence 10697, A
23	31	55.4	336	9	US-09-782-974C-86	Sequence 86, Appl
24	31	55.4	337	9	US-10-023-775B-2	Sequence 2, Appli
25	31	55.4	337	9	US-10-027-144-2	Sequence 2, Appli
26	31	55.4	337	9	US-10-188-405-8	Sequence 8, Appli
27	31	55.4	337	9	US-09-885-453-1	Sequence 1, Appli
28	31	55.4	337	10	US-09-943-798-4	Sequence 4, Appli
29	31	55.4	447	10	US-09-765-272-182	Sequence 182, App
30	31	55.4	449	9	US-09-738-626-5315	Sequence 5315, Ap
31	31	55.4	484	9	US-09-769-787-38	Sequence 38, Appl
32	31	55.4	529	10	US-09-923-304-4	Sequence 4, Appli
33	31	55.4	763	10	US-09-765-272-66	Sequence 66, Appl
34	31	55.4	796	10	US-09-765-272-56	Sequence 56, Appl
35	31	55.4	826	9	US-09-769-787-194	Sequence 194, App
36	31	55.4	838	9	US-09-884-465A-8	Sequence 8, Appli
37	31	55.4	840	9	US-09-884-465A-7	Sequence 7, Appli
38	31	55.4	840	9	US-09-884-465A-10	Sequence 10, Appl
39	31	55.4	1039	9	US-09-884-465A-6	Sequence 6, Appli
40	31	55.4	1463	9	US-10-176-847-22	Sequence 22, Appl
41	31	55.4	1724	9	US-09-964-899-43	Sequence 43, Appl
42	30	53.6	7	9	US-09-909-062-1	Sequence 1, Appli
43	30	53.6	7	9	US-09-909-062-9	Sequence 9, Appli
44	30	53.6	7	9	US-09-909-062-130	Sequence 130, App
45	30	53.6	44	10	US-09-881-752A-292	Sequence 292, App
46	30	53.6	72	10	US-09-925-297-664	Sequence 664, App
47	30	53.6	121	9	US-09-852-797-68	Sequence 68, Appl
48	30	53.6	121	9	US-09-852-797-85	Sequence 85, Appl
49	30	53.6	121	10	US-09-853-161-68	Sequence 68, Appl
50	30	53.6	121	10	US-09-853-161-85	Sequence 85, Appl
51	30	53.6	121	10	US-09-852-659A-68	Sequence 68, Appl
52	30	53.6	121	10	US-09-852-659A-85	Sequence 85, Appl
53	30	53.6	135	9	US-09-992-598-359	Sequence 359, App
54	30	53.6	135	9	US-09-989-293A-359	Sequence 359, App
55	30	53.6	135	9	US-09-989-735-359	Sequence 359, App
56	30	53.6	135	9	US-09-989-444-359	Sequence 359, App
57	30	53.6	135	9	US-09-989-730-359	Sequence 359, App
58	30	53.6	135	9	US-09-989-436-359	Sequence 359, App
59	30	53.6	135	9	US-09-991-181-359	Sequence 359, App
60	30	53.6	135	9	US-09-993-687-359	Sequence 359, App
61	30	53.6	135	9	US-09-989-734-359	Sequence 359, App
62	30	53.6	135	9	US-09-997-653-359	Sequence 359, App
63	30	53.6	135	9	US-10-174-590-444	Sequence 444, App
64	30	53.6	135	9	US-10-176-758-444	Sequence 444, App
65	30	53.6	135	9	US-10-175-737-444	Sequence 444, App
66	30	53.6	135	9	US-09-993-667-359	Sequence 359, App
67	30	53.6	135	9	US-10-173-706-444	Sequence 444, App
68	30	53.6	135	9	US-10-175-738-444	Sequence 444, App
69	30	53.6	135	9	US-10-175-752-444	Sequence 444, App
70	30	53.6	135	9	US-10-176-482-444	Sequence 444, App
71	30	53.6	135	9	US-10-176-757-444	Sequence 444, App
72	30	53.6	135	9	US-10-176-913-444	Sequence 444, App
73	30	53.6	135	9	US-10-180-552-444	Sequence 444, App
74	30	53.6	135	9	US-10-180-557-444	Sequence 444, App
75	30	53.6	135	9	US-09-990-438-359	Sequence 359, App

```

; PRIOR APPLICATION NUMBER: 60/168,677
; PRIOR FILING DATE: 1999-12-03
; PRIOR APPLICATION NUMBER: 60/175,691
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/191,638
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 198
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 126
; LENGTH: 567
; TYPE: PRT
; ORGANISM: Drosophila
US-10-270-333-126

```

```

Query Match      58.9%; Score 33; DB 9; Length 567;
Best Local Similarity 83.3%; Pred. No. 2.8e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY      5 PXGMHY 10
      | | | | |
Db      402 PGMHY 407

```

## RESULT 2

```

US-09-738-626-4919
; Sequence 4919, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIO
; APPLICANT: OCHIALI, KEIKO
; APPLICANT: YOKOL, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENO, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 4919
; LENGTH: 622
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-4919

```

```

Query Match      58.9%; Score 33; DB 9; Length 622;
Best Local Similarity 50.0%; Pred. No. 3.1e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 2; Gaps 1;

```

```

QY      1 EEVVPXGM--HY 10
      | | | | |
Db      46 EEIIPGVPNHY 57

```

## RESULT 3

```

US-09-866-050A-663
; Sequence 663, Application US/09866050A
; Publication No. US20030040471A1
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew

```

```

; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James G.
; APPLICANT: Kumble, Krishanand D.
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; TITLE OF INVENTION: and Methods for Their Use
; FILE REFERENCE: 11000.1011c4U
; CURRENT APPLICATION NUMBER: US/09/866,050A
; CURRENT FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 725
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 663
; LENGTH: 747
; TYPE: PRT
; ORGANISM: Rat
US-09-866-050A-663

```

```

Query Match      58.9%; Score 33; DB 9; Length 747;
Best Local Similarity 71.4%; Pred. No. 3.8e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

QY      5 PXGMHYS 11
      | | | | |
Db      627 PGLHYS 633

```

## RESULT 4

```

US-10-027-806-4
; Sequence 4, Application US/10027806
; Patent No. US20020160476A1
; GENERAL INFORMATION:
; APPLICANT: Swanson, Ronald V.
; APPLICANT: Feldman, Robert A.
; APPLICANT: Schleper, Christa
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM CENARCHAEUM SYMBIOSUM
; FILE REFERENCE: DCCRP.002A
; CURRENT APPLICATION NUMBER: US/10/027,806
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/408,020
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-09-29
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 3472
; TYPE: PRT
; ORGANISM: Cenarchaeum symbiosum
US-10-027-806-4

```

```

Query Match      58.9%; Score 33; DB 9; Length 3472;
Best Local Similarity 45.5%; Pred. No. 2.1e+03;
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

```

```

QY      1 EEVVPXGMHYS 11
      | | | | |
Db      2294 EDVIPRGISFS 2304

```

## RESULT 5

```

US-10-034-623-4
; Sequence 4, Application US/10034623
; Publication No. US20020198365A1
; GENERAL INFORMATION:
; APPLICANT: Swanson, Ronald V.
; APPLICANT: Feldman, Robert A.
; APPLICANT: Schleper, Christa
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM CENARCHAEUM SYMBIOSUM
; FILE REFERENCE: DCCRP.002A
; CURRENT APPLICATION NUMBER: US/10/034,623
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 09/408,020
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: 60/102,294
; PRIOR FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 123

```

; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 4  
; LENGTH: 3472  
; TYPE: PRT  
; ORGANISM: Cenarchaeum symbiosum  
US-10-034-623-4

Query Match 57.1%; Score 32; DB 10; Length 254;  
Best Local Similarity 60.0%; Pred. No. 1.8e+02;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMHY 10  
||| | :|  
DB 74 EEVFPAMNY 83

RESULT 8  
US-10-024-066-2  
; Sequence 2, Application US/10024066  
; Patent No. US20020166134A1  
; GENERAL INFORMATION:  
; APPLICANT: Field, Loren J.  
; APPLICANT: Pasumarthi, Kishore Babu S.  
; TITLE OF INVENTION: CARDIOMYOCYTES WITH ENHANCED PROLIFERATIVE POTENTIAL,  
; TITLE OF INVENTION: AND METHODS FOR PREPARING AND USING SAME  
; FILE REFERENCE: 7037-450  
; CURRENT APPLICATION NUMBER: US/10/024,066  
; CURRENT FILING DATE: 2001-12-18  
; PRIOR APPLICATION NUMBER: 60/139,942  
; PRIOR FILING DATE: 1999-06-18  
; PRIOR APPLICATION NUMBER: PCT/US00/16827  
; PRIOR FILING DATE: 2000-06-19  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 289  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-10-024-066-2

Query Match 57.1%; Score 32; DB 9; Length 289;  
Best Local Similarity 60.0%; Pred. No. 2.1e+02;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMHY 10  
||| | :|  
DB 73 EEVFPAMNY 82

RESULT 9  
US-10-024-066-4  
; Sequence 4, Application US/10024066  
; Patent No. US20020166134A1  
; GENERAL INFORMATION:  
; APPLICANT: Field, Loren J.  
; APPLICANT: Pasumarthi, Kishore Babu S.  
; TITLE OF INVENTION: CARDIOMYOCYTES WITH ENHANCED PROLIFERATIVE POTENTIAL,  
; TITLE OF INVENTION: AND METHODS FOR PREPARING AND USING SAME  
; FILE REFERENCE: 7037-450  
; CURRENT APPLICATION NUMBER: US/10/024,066  
; CURRENT FILING DATE: 2001-12-18  
; PRIOR APPLICATION NUMBER: 60/139,942  
; PRIOR FILING DATE: 1999-06-18  
; PRIOR APPLICATION NUMBER: PCT/US00/16827  
; PRIOR FILING DATE: 2000-06-19  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 289  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-024-066-4

Query Match 57.1%; Score 32; DB 9; Length 289;  
Best Local Similarity 60.0%; Pred. No. 2.1e+02;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMHY 10  
||| | :|

; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 4  
; LENGTH: 3472  
; TYPE: PRT  
; ORGANISM: Cenarchaeum symbiosum  
US-10-034-623-4

Query Match 58.9%; Score 33; DB 9; Length 3472;  
Best Local Similarity 45.5%; Pred. No. 2.1e+03;  
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMHYS 11  
|:|:| | :|  
DB 2294 EDVIPRGISFS 2304

RESULT 6  
US-10-027-801-4  
; Sequence 4, Application US/10027801  
; Publication No. US20030054364A1  
; GENERAL INFORMATION:  
; APPLICANT: Swanson, Ronald V.  
; APPLICANT: Feldman, Robert A.  
; APPLICANT: Schleper, Christa  
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM CENARCHAEUM SYMBIOSUM  
; FILE REFERENCE: DCCP.002A  
; CURRENT APPLICATION NUMBER: US/10/027,801  
; CURRENT FILING DATE: 2001-12-21  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/408,020  
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-09-29  
; NUMBER OF SEQ ID NOS: 123  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 4  
; LENGTH: 3472  
; TYPE: PRT  
; ORGANISM: Cenarchaeum symbiosum  
US-10-027-801-4

Query Match 58.9%; Score 33; DB 9; Length 3472;  
Best Local Similarity 45.5%; Pred. No. 2.1e+03;  
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMHYS 11  
|:|:| | :|  
DB 2294 EDVIPRGISFS 2304

RESULT 7  
US-09-778-927A-53  
; Sequence 53, Application US/09778927A  
; Patent No. US20020068342A1  
; GENERAL INFORMATION:  
; APPLICANT: KHOSRAVI, Rami et al.  
; TITLE OF INVENTION: NOVEL NUCLEIC ACID AND AMINO ACID SEQUENCES AND NOVEL  
; TITLE OF INVENTION: VARIANTS OF ALTERNATIVE SPLICING  
; FILE REFERENCE: 2786-0160P  
; CURRENT APPLICATION NUMBER: US/09/778,927A  
; CURRENT FILING DATE: 2001-02-08  
; PRIOR APPLICATION NUMBER: IL 134453  
; PRIOR FILING DATE: 2000-02-09  
; PRIOR APPLICATION NUMBER: IL135341  
; PRIOR FILING DATE: 2000-03-29  
; NUMBER OF SEQ ID NOS: 81  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 53  
; LENGTH: 254  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)..(254 )  
; OTHER INFORMATION: Xaa = any amino acid, unknown, or other  
US-09-778-927A-53

Db 74 EEVFPPLAMNY 83

## RESULT 10

US-09-919-497-54  
; Sequence 54, Application US/09919497  
; Patent No. US20020106662A1  
; GENERAL INFORMATION:  
; APPLICANT: Mutter, George L.  
; TITLE OF INVENTION: PROGNOSTIC CLASSIFICATION OF ENDOMETRIAL CANCER  
; FILE REFERENCE: B0801/7225  
; CURRENT APPLICATION NUMBER: US/09/919,497  
; PRIOR FILING DATE: 2001-07-31  
; PRIOR APPLICATION NUMBER: US 60/221,735  
; PRIOR FILING DATE: 2000-07-31  
; NUMBER OF SEQ ID NOS: 100  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 54  
; LENGTH: 289  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-919-497-54

Query Match 57.1%; Score 32; DB 10; Length 289;  
Best Local Similarity 60.0%; Pred. No. 2.1e+02;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMHY 10

Db 74 EEVFPPLAMNY 83

## RESULT 11

US-09-820-843A-26  
; Sequence 26, Application US/09820843A  
; Publication No. US20030039963A1  
; GENERAL INFORMATION:  
; APPLICANT: Council of Scientific and Industrial Research  
; TITLE OF INVENTION: A COMPUTATIONAL METHOD FOR THE IDENTIFICATION OF CANDIDATE PROTEIN  
; FILE REFERENCE: 063915  
; CURRENT APPLICATION NUMBER: US/09/820,843A  
; CURRENT FILING DATE: 2001-03-30  
; NUMBER OF SEQ ID NOS: 118  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 26  
; LENGTH: 653  
; TYPE: PRT  
; ORGANISM: Vibrio cholerae  
; FEATURE:  
; NAME/KEY: misc.feature  
; OTHER INFORMATION: iron(III) ABC transporter, permease protein  
; NAME/KEY: misc.feature  
; OTHER INFORMATION: gi|9654609  
US-09-820-843A-26

Query Match 57.1%; Score 32; DB 9; Length 653;  
Best Local Similarity 75.0%; Pred. No. 5.1e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXGM 8

Db 300 EEVPSGI 307

## RESULT 12

US-09-252-088-16  
; Sequence 16, Application US/09252088  
; Publication No. US20030031682A1  
; GENERAL INFORMATION:  
; APPLICANT: BRODEUR, Bernard R.  
; APPLICANT: RIOUX, Clment  
; APPLICANT: BOYER, Martine

; APPLICANT: CHARLEBOIS, Isabelle  
; APPLICANT: HAMEL, Jose  
; APPLICANT: MARTIN, Denis  
; TITLE OF INVENTION: NOVEL GROUP B STREPTOCOCCUS ANTIGENS  
; FILE REFERENCE: 8331-9002  
; CURRENT APPLICATION NUMBER: US/09/252,088  
; CURRENT FILING DATE: 1999-02-18  
; EARLIER APPLICATION NUMBER: US/60/075,425  
; EARLIER FILING DATE: 1998-02-20  
; NUMBER OF SEQ ID NOS: 44  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 16  
; LENGTH: 715  
; TYPE: PRT  
; ORGANISM: group B streptococcus  
US-09-252-088-16

Query Match 57.1%; Score 32; DB 9; Length 715;  
Best Local Similarity 75.0%; Pred. No. 5.7e+02;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 VVPXGMHY 10

Db 243 VVPHGDHY 250

## RESULT 13

US-09-252-088-15  
; Sequence 15, Application US/09252088  
; Publication No. US20030031682A1  
; GENERAL INFORMATION:  
; APPLICANT: BRODEUR, Bernard R.  
; APPLICANT: BOYER, Martine  
; APPLICANT: CHARLEBOIS, Isabelle  
; APPLICANT: HAMEL, Jose  
; APPLICANT: MARTIN, Denis  
; TITLE OF INVENTION: NOVEL GROUP B STREPTOCOCCUS ANTIGENS  
; FILE REFERENCE: 8331-9002  
; CURRENT APPLICATION NUMBER: US/09/252,088  
; CURRENT FILING DATE: 1999-02-18  
; EARLIER APPLICATION NUMBER: US/60/075,425  
; EARLIER FILING DATE: 1998-02-20  
; NUMBER OF SEQ ID NOS: 44  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 15  
; LENGTH: 793  
; TYPE: PRT  
; ORGANISM: group B streptococcus  
US-09-252-088-15

Query Match 57.1%; Score 32; DB 9; Length 793;  
Best Local Similarity 75.0%; Pred. No. 6.4e+02;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 VVPXGMHY 10

Db 321 VVPHGDHY 328

## RESULT 14

US-10-047-542-80  
; Sequence 80, Application US/10047542  
; Patent No. US20020168367A1  
; GENERAL INFORMATION:  
; APPLICANT: LARRICK, JAMES W.  
; APPLICANT: WYCOFF, KEITH L.  
; TITLE OF INVENTION: NOVEL IMMUNOADHESINS FOR TREATING AND PREVENTING VIRAL  
; FILE REFERENCE: 030905.0004.C1P1  
; CURRENT APPLICATION NUMBER: US/10/047,542  
; CURRENT FILING DATE: 2001-10-26  
; PRIOR APPLICATION NUMBER: PCT/US01/13932



; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.2  
 ; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.8  
 ; OTHER INFORMATION: EST\_HUMAN HIT: AWO27739.1, EVALUE 7.00e-14  
 US-09-864-761-40832

Query Match 55.4%; Score 31; DB 10; Length 73;  
 Best Local Similarity 71.4%; Pred. No. 72;  
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMH 9  
 Db 60 VVPXGLH 66  
 ||| |||

RESULT 18  
 US-10-001-857-119  
 ; Sequence 119, Application US/10001857  
 ; Publication No. US20020183500A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Macina, Roberto  
 ; APPLICANT: Recipon, Herve  
 ; APPLICANT: Chen, Sei-yu  
 ; APPLICANT: Sun, Yongming  
 ; APPLICANT: Liu, Chenghua  
 ; TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Proteins  
 ; FILE REFERENCE: DEX-0273  
 ; CURRENT APPLICATION NUMBER: US/10/001,857  
 ; CURRENT FILING DATE: 2001-11-20  
 ; PRIOR APPLICATION NUMBER: 60/252,054  
 ; PRIOR FILING DATE: 2000-11-20  
 ; NUMBER OF SEQ ID NOS: 208  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 119  
 ; LENGTH: 192  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapien  
 US-10-001-857-119

Query Match 55.4%; Score 31; DB 9; Length 192;  
 Best Local Similarity 57.1%; Pred. No. 2.1e+02;  
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMH 9  
 Db 145 IIPKGMH 151  
 :|| |||

RESULT 19  
 US-09-986-480-171  
 ; Sequence 171, Application US/09986480  
 ; Publication No. US20030027999A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Rosen et al.  
 ; TITLE OF INVENTION: 143 Human Secreted Proteins  
 ; FILE REFERENCE: PS500P1  
 ; CURRENT APPLICATION NUMBER: US/09/986,480  
 ; CURRENT FILING DATE: 2001-11-08  
 ; PRIOR APPLICATION NUMBER: PCT/US00/12788  
 ; PRIOR FILING DATE: 2000-05-11  
 ; PRIOR APPLICATION NUMBER: US 60/134,068  
 ; PRIOR FILING DATE: 1998-05-13  
 ; NUMBER OF SEQ ID NOS: 456  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 171  
 ; LENGTH: 192  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-986-480-171

Query Match 55.4%; Score 31; DB 9; Length 192;  
 Best Local Similarity 50.0%; Pred. No. 2.1e+02;  
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 EVVPXGMHYS 11  
 Db 52 ENIPEGLNYS 61  
 | : | | : | |

RESULT 20  
 US-09-815-242-13489  
 ; Sequence 13489, Application US/09815242  
 ; Patent No. US20020061569A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Haselbeck, Robert  
 ; APPLICANT: Ohlsen, Kari L.  
 ; APPLICANT: Zyskind, Judith W.  
 ; APPLICANT: Wall, Daniel  
 ; APPLICANT: Trawick, John D.  
 ; APPLICANT: Carr, Grant J.  
 ; APPLICANT: Yamamoto, Robert T.  
 ; APPLICANT: Xu, H. Howard  
 ; TITLE OF INVENTION: Identification of Essential Genes in  
 ; TITLE OF INVENTION: Prokaryotes  
 ; FILE REFERENCE: ELITRA.011A  
 ; CURRENT APPLICATION NUMBER: US/09/815,242  
 ; CURRENT FILING DATE: 2001-03-21  
 ; PRIOR APPLICATION NUMBER: 60/191,078  
 ; PRIOR FILING DATE: 2000-03-21  
 ; PRIOR APPLICATION NUMBER: 60/206,848  
 ; PRIOR FILING DATE: 2000-05-23  
 ; PRIOR APPLICATION NUMBER: 60/207,727  
 ; PRIOR FILING DATE: 2000-05-26  
 ; PRIOR APPLICATION NUMBER: 60/242,578  
 ; PRIOR FILING DATE: 2000-10-23  
 ; PRIOR APPLICATION NUMBER: 60/253,625  
 ; PRIOR FILING DATE: 2000-11-27  
 ; PRIOR APPLICATION NUMBER: 60/257,931  
 ; PRIOR FILING DATE: 2000-12-22  
 ; PRIOR APPLICATION NUMBER: 60/269,308  
 ; PRIOR FILING DATE: 2001-02-16  
 ; NUMBER OF SEQ ID NOS: 14110  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 13489  
 ; LENGTH: 260  
 ; TYPE: PRT  
 ; ORGANISM: Streptococcus pneumoniae  
 US-09-815-242-13489

Query Match 55.4%; Score 31; DB 10; Length 260;  
 Best Local Similarity 60.0%; Pred. No. 2.9e+02;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMHY 10  
 Db 188 EEKVGEGVHY 197  
 ||| |||

RESULT 21  
 US-09-815-242-13613  
 ; Sequence 13613, Application US/09815242  
 ; Patent No. US20020061569A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Haselbeck, Robert  
 ; APPLICANT: Ohlsen, Kari L.  
 ; APPLICANT: Zyskind, Judith W.  
 ; APPLICANT: Wall, Daniel  
 ; APPLICANT: Trawick, John D.  
 ; APPLICANT: Carr, Grant J.  
 ; APPLICANT: Yamamoto, Robert T.  
 ; APPLICANT: Xu, H. Howard  
 ; TITLE OF INVENTION: Identification of Essential Genes in  
 ; TITLE OF INVENTION: Prokaryotes  
 ; FILE REFERENCE: ELITRA.011A  
 ; CURRENT APPLICATION NUMBER: US/09/815,242  
 ; CURRENT FILING DATE: 2001-03-21  
 ; PRIOR APPLICATION NUMBER: 60/191,078



;; PRIOR FILING DATE: 2000-03-21  
;; PRIOR APPLICATION NUMBER: 60/206,848  
;; PRIOR FILING DATE: 2000-05-23  
;; PRIOR APPLICATION NUMBER: 60/207,727  
;; PRIOR FILING DATE: 2000-05-26  
;; PRIOR APPLICATION NUMBER: 60/242,578  
;; PRIOR FILING DATE: 2000-10-23  
;; PRIOR APPLICATION NUMBER: 60/253,625  
;; PRIOR FILING DATE: 2000-11-27  
;; PRIOR APPLICATION NUMBER: 60/257,931  
;; PRIOR FILING DATE: 2000-12-22  
;; PRIOR APPLICATION NUMBER: 60/269,308  
;; PRIOR FILING DATE: 2001-02-16  
;; NUMBER OF SEQ ID NOS: 14110  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 13613  
;; LENGTH: 260  
;; TYPE: PRT  
;; ORGANISM: Streptococcus pneumoniae  
US-09-815-242-13613

Query Match 55.4%; Score 31; DB 10; Length 260;  
Best Local Similarity 60.0%; Pred. No. 2.9e+02;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMHY 10  
||| |:  
DB 188 EEKVGGVHY 197

## RESULT 22

;; Sequence 10697, Application US/09815242  
;; Patent No. US2002061569A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Haselbeck, Robert  
;; APPLICANT: Ohlsen, Karl L.  
;; APPLICANT: Zyskind, Judith W.  
;; APPLICANT: Wall, Daniel  
;; APPLICANT: Trawick, John D.  
;; APPLICANT: Carr, Grant J.  
;; APPLICANT: Yamamoto, Robert T.  
;; APPLICANT: Xu, H. Howard  
;; TITLE OF INVENTION: Identification of Essential Genes in  
;; TITLE OF INVENTION: Prokaryotes  
;; FILE REFERENCE: ELITRA.011a  
;; CURRENT APPLICATION NUMBER: US/09/815,242  
;; PRIOR FILING DATE: 2001-03-21  
;; PRIOR APPLICATION NUMBER: 60/191,078  
;; PRIOR FILING DATE: 2000-03-21  
;; PRIOR APPLICATION NUMBER: 60/206,848  
;; PRIOR FILING DATE: 2000-05-23  
;; PRIOR APPLICATION NUMBER: 60/207,727  
;; PRIOR FILING DATE: 2000-05-26  
;; PRIOR APPLICATION NUMBER: 60/242,578  
;; PRIOR FILING DATE: 2000-10-23  
;; PRIOR APPLICATION NUMBER: 60/253,625  
;; PRIOR FILING DATE: 2000-11-27  
;; PRIOR APPLICATION NUMBER: 60/257,931  
;; PRIOR FILING DATE: 2000-12-22  
;; PRIOR APPLICATION NUMBER: 60/269,308  
;; PRIOR FILING DATE: 2001-02-16  
;; NUMBER OF SEQ ID NOS: 14110  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 10697  
;; LENGTH: 299  
;; TYPE: PRT  
;; ORGANISM: Enterococcus faecalis  
US-09-815-242-10697

Query Match 55.4%; Score 31; DB 10; Length 299;  
Best Local Similarity 40.0%; Pred. No. 3.4e+02;  
Matches 4; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMHY 10  
|:: | |:  
DB 218 EQITPTGIEY 227

## RESULT 23

US-09-782-974C-86  
;; Sequence 86, Application US/09782974C  
;; Publication No. US20030082534A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Vogeli, Gabriel  
;; APPLICANT: Lind, Peter  
;; APPLICANT: Wood, Linda S.  
;; APPLICANT: Parodi, Luis A.  
;; TITLE OF INVENTION: No. US20030082534A1 G Protein Coupled Receptor  
;; FILE REFERENCE: A1IUSPHRM311  
;; CURRENT APPLICATION NUMBER: US/09/782,974C  
;; CURRENT FILING DATE: 2002-09-04  
;; PRIOR APPLICATION NUMBER: 60/165,838  
;; PRIOR FILING DATE: 1999-11-16  
;; PRIOR APPLICATION NUMBER: 09/714,449  
;; PRIOR FILING DATE: 2000-11-16  
;; PRIOR APPLICATION NUMBER: 60/198,568  
;; PRIOR FILING DATE: 2000-04-20  
;; PRIOR APPLICATION NUMBER: 60/166,071  
;; PRIOR FILING DATE: 1999-11-17  
;; PRIOR APPLICATION NUMBER: 60/166,678  
;; PRIOR FILING DATE: 1999-11-19  
;; PRIOR APPLICATION NUMBER: 60/173,396  
;; PRIOR FILING DATE: 1999-12-28  
;; PRIOR APPLICATION NUMBER: 60/184,129  
;; PRIOR FILING DATE: 2000-02-22  
;; PRIOR APPLICATION NUMBER: 60/185,421  
;; PRIOR FILING DATE: 2000-02-28  
;; PRIOR APPLICATION NUMBER: 60/185,554  
;; PRIOR FILING DATE: 2000-02-28  
;; PRIOR APPLICATION NUMBER: 60/186,530  
;; PRIOR FILING DATE: 2000-03-02  
;; Remaining Prior Application data removed - See File Wrapper or PALM.  
;; NUMBER OF SEQ ID NOS: 192  
;; SOFTWARE: PatentIn version 3.1  
;; SEQ ID NO 86  
;; LENGTH: 336  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-09-782-974C-86

Query Match 55.4%; Score 31; DB 9; Length 336;  
Best Local Similarity 50.0%; Pred. No. 3.8e+02;  
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMHY 10  
:|: | |:  
DB 26 DENIPLKMHY 35

## RESULT 24

US-10-023-775B-2  
;; Sequence 2, Application US/10023775B  
;; Publication No. US20030022828A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Pfizer Inc. (US, JP, EB except GB)  
;; APPLICANT: Fidock, Mark David  
;; TITLE OF INVENTION: No. US20030022828A1 Polypeptide  
;; FILE REFERENCE: PCI0959AGPR  
;; CURRENT APPLICATION NUMBER: US/10/023,775B  
;; CURRENT FILING DATE: 2001-12-18  
;; PRIOR APPLICATION NUMBER: GB 0030854.4  
;; PRIOR FILING DATE: 2000-12-18  
;; PRIOR APPLICATION NUMBER: US 60/260,590  
;; PRIOR FILING DATE: 2001-01-09

; PRIOR APPLICATION NUMBER: US 60/296,660  
; PRIOR FILING DATE: 2001-06-07  
; PRIOR APPLICATION NUMBER: GB 0111031.1  
; PRIOR FILING DATE: 2001-05-04  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2  
; LENGTH: 337  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-023-775B-2

Query Match 55.4%; Score 31; DB 9; Length 337;  
Best Local Similarity 50.0%; Pred. No. 3.8e+02;  
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMHY 10  
:|:|  
Db 26 DENIPLKMHY 35

## RESULT 25

US-10-270-144-2  
; Sequence 2, Application US/10270144  
; Publication No. US20030049790A1  
; GENERAL INFORMATION:  
; APPLICANT: WEI, Ming-Hui et al  
; TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED  
; TITLE OF INVENTION: RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR  
; FILE REFERENCE: CL000750CON  
; CURRENT APPLICATION NUMBER: US/10/270,144  
; CURRENT FILING DATE: 2002-10-15  
; PRIOR APPLICATION NUMBER: 60/205,196  
; PRIOR FILING DATE: 2000-05-18  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 337  
; TYPE: PRT  
; ORGANISM: Human  
US-10-270-144-2

Query Match 55.4%; Score 31; DB 9; Length 337;  
Best Local Similarity 50.0%; Pred. No. 3.8e+02;  
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMHY 10  
:|:|  
Db 26 DENIPLKMHY 35

Search completed: June 10, 2003, 14:35:43  
Job time : 15.0714 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2003, 13:31:15 ; Search time 11.2143 Seconds  
(without alignments)  
94.297 Million cell updates/sec

Title: US-09-909-164-11  
Perfect score: 56  
Sequence: 1 EEVVPXGMHYS 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 75 summaries

Database : PIR\_73:\*

1: pir1:\*\*\*  
2: pir2:\*\*\*  
3: pir3:\*\*\*  
4: pir4:\*\*\*

pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	37	66.1	308	A72207	ftsH proteinase ac
2	37	66.1	1057	F89892	carbamoyl-phosphat
3	36	64.3	102	A42452	VI protein - tobac
4	36	64.3	252	AE2001	hypothetical prote
5	36	64.3	460	S69046	hypothetical prote
6	36	64.3	743	S38143	hypothetical prote
7	35	62.5	156	D82618	conserved hypothet
8	35	62.5	233	T02590	DNA binding protein
9	35	62.5	311	H69194	GMP synthetase, su
10	35	62.5	425	T24111	hypothetical prote
11	35	62.5	510	G86430	T518.1 protein - A
12	34	60.7	264	C69117	diphthine synthase
13	34	60.7	279	C75558	hypothetical prote
14	34	60.7	350	B75478	3-dehydroquinatase
15	34	60.7	355	T35025	probable DNA ligas
16	34	60.7	360	E69086	cell division prot
17	34	60.7	425	C83903	hypothetical prote
18	34	60.7	425	S58132	hypothetical prote
19	34	60.7	495	T28717	ATP-dependent DNA
20	34	60.7	1028	AF3286	hypothetical prote
21	33	58.9	156	S54619	polyamine transpor
22	33	58.9	367	E83607	conserved hypothet
23	33	58.9	441	G82253	probable amino aci
24	33	58.9	466	G71542	amino acid antipor
25	33	58.9	466	H81697	finger protein (cl
26	33	58.9	487	S65811	cytochrome-c3 hydr
27	33	58.9	514	1 HQDVLB	coenzyme F420-quin
28	33	58.9	534	A69284	hypothetical prote
29	33	58.9	545	T08564	

30	33	58.9	627	2	A69663	DNA mismatch repai
31	33	58.9	716	1	JC5061	macrophage-stimula
32	33	58.9	1257	2	S44754	C14B9, 8 protein -
33	33	58.9	1396	2	S38851	L-shaped tail fibe
34	33	58.9	3472	2	T31308	hypothetical 367K
35	32	57.1	126	2	A25905	tachykinin B precu
36	32	57.1	197	2	D71640	heme exporter prot
37	32	57.1	225	2	S57810	hypothetical prote
38	32	57.1	233	2	E97120	ribosomal protein
39	32	57.1	267	2	T07215	probable Thua prot
40	32	57.1	270	2	C95881	hypothetical prote
41	32	57.1	283	2	T25737	hypothetical prote
42	32	57.1	288	2	JC4011	cyclin D2 - rat
43	32	57.1	288	2	I58372	cyclin D2 - mouse
44	32	57.1	289	2	A41984	cyclin D2 - human
45	32	57.1	289	2	A42822	cyclin D1 - Africa
46	32	57.1	291	2	S57925	cyclin D2 - Africa
47	32	57.1	291	2	JC4579	cyclin D2 - chicke
48	32	57.1	291	2	S62730	cyclin D1 - zebra
49	32	57.1	292	2	B42822	cyclin D3 - human
50	32	57.1	292	2	A38977	cyclin D1 - human
51	32	57.1	295	2	A56523	cyclin D1 - mouse
52	32	57.1	295	2	JC2342	cyclin D1 - rat
53	32	57.1	295	2	T25498	hypothetical prote
54	32	57.1	335	2	S43354	hypothetical prote
55	32	57.1	341	2	JX0116	type II site-speci
56	32	57.1	354	2	A83470	probable periplasm
57	32	57.1	363	2	T17267	hypothetical prote
58	32	57.1	369	2	G69119	probable aspartate
59	32	57.1	374	2	AD0049	hypothetical prote
60	32	57.1	384	2	E82626	hypothetical prote
61	32	57.1	405	2	A48702	2-methyl-branched-
62	32	57.1	412	2	S51348	hypothetical prote
63	32	57.1	509	2	G84598	probable bZIP tran
64	32	57.1	519	2	F69263	conserved hypothet
65	32	57.1	534	2	AC1421	beta-glucoside-spe
66	32	57.1	617	2	D82352	iron(III) ABC tran
67	32	57.1	653	2	S22293	zinc finger protei
68	32	57.1	670	2	A91247	phage transposase
69	32	57.1	696	2	T29996	hypothetical prote
70	32	57.1	704	2	S25330	SCRI protein - yea
71	32	57.1	759	2	T46758	hypothetical 92.4K
72	32	57.1	822	2	T39116	probable sulfate p
73	32	57.1	840	2	T40413	sulfate permease -
74	32	57.1	877	2	T03104	tegument protein h
75	32	57.1	1369	2		

#### ALIGNMENTS

#### RESULT 1

A72207

ftsH proteinase activity modulator HflK - Thermotoga maritima (strain MSB8)

C:Species: Thermotoga maritima

C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 21-Jul-2000

C:Accession: A72207

R:Neilson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.;

Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson

C.M.

Nature 399, 323-329, 1999

A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genom

A:Reference number: A72200; MUID:99287316; PMID:10360571

A:Accession: A72207

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-308 <ARN>

A:Cross-references: GB:AE001819; GB:AE000512; NID:g4982396; PIDN:AAD36885.1; PID:g

A:Experimental source: strain MSB8

C:Genetics:

A:Gene: TWI822

C:Superfamily: erythrocyte band 7 integral membrane protein

Query Match 66.1%; Score 37; DB 2; Length 308;  
 Best Local Similarity 75.0%; Pred. No. 10;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMHY 10  
 ||| |||  
 Db 41 VVPSGIHY 48

## RESULT 2

F89892 carbamoyl-phosphate synthase large chain [imported] - Staphylococcus aureus (strain N315)

C:Species: Staphylococcus aureus  
 C>Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 22-Oct-2001  
 C:Accession: F89892

R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogud  
 ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;  
 C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.  
 Lancet 357, 1225-1240, 2001

A:Title: Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*.

A:Reference number: A89758; MUID:21311952; PMID:11418146

A:Accession: F89892

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1057 <KUR>

A:CROSS-references: GB:BA000018; PID:g13701002; PIDN:BA042298.1; GSPDB:GN00149

A:Experimental source: strain N315

C:Genetics:

C:Superfamily: carbamoyl-phosphate synthase (glutamine-hydrolyzing) large chain; biotin

Query Match 66.1%; Score 37; DB 2; Length 1057;  
 Best Local Similarity 60.0%; Pred. No. 39;  
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 EVVPXGMHYS 11  
 ||| |||  
 Db 190 EIVSNGLHYS 199

## RESULT 3

A42452 V1 protein - tobacco yellow dwarf virus (strain Australia)

C:Species: tobacco yellow dwarf virus

C>Date: 15-Jan-1993 #sequence\_revision 15-Jan-1993 #text\_change 08-Oct-1999

C:Accession: A42452

R:Morris, B.A.M.; Richardson, K.A.; Haley, A.; Zhan, X.; Thomas, J.E.

Virology 187, 633-642, 1992

A:Title: The nucleotide sequence of the infectious cloned DNA component of tobacco yellow

A:Reference number: A42452; MUID:921188538; PMID:1546458

A:Accession: A42452

A:Molecule type: DNA

A:Residues: 1-102 <MOR>

A:CROSS-references: GB:M81103; NID:g335283; PIDN:AAA47947.1; PID:g335284

Query Match 64.3%; Score 36; DB 2; Length 102;  
 Best Local Similarity 60.0%; Pred. No. 5;  
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 EVVPXGMHYS 11  
 ||| |||  
 Db 7 QVVPSGINYS 16

## RESULT 4

AZ0001 hypothetical protein alr1563 [imported] - Nostoc sp. (strain PCC 7120)

C:Species: Nostoc sp.

A:Note: Nostoc sp. strain PCC 7120 is a synonym of *Anabaena* sp. strain PCC 7120

C>Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 30-Jun-2002

C:Accession: AE2001

R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriiguchi

Nakazaki, N.; Shimpou, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S

DNA Res. 8, 205-213, 2001

A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing *Cyanobacterium*

A:Reference number: AB1807; MUID:21595285; PMID:111759840

A:Accession: AE2001

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-252 <KUR>

A:CROSS-references: GB:BA000019; PIDN:BA077929.1; PID:g17135383; GSPDB:GN00179

A:Experimental source: strain PCC 7120

C:Genetics:

A:Gene: alr1563

Query Match 64.3%; Score 36; DB 2; Length 252;  
 Best Local Similarity 50.0%; Pred. No. 13;  
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMHY 10  
 | || | : | : |  
 Db 235 EMIVPAGLHF 244

## RESULT 5

S69046

hypothetical protein YPL139c - yeast (*Saccharomyces cerevisiae*)

C:Species: *Saccharomyces cerevisiae*

C>Date: 22-Aug-1996 #sequence\_revision 06-Sep-1996 #text\_change 17-Mar-2000

C:Accession: S69046

R:Hall, J.; DePaulo, T.; Ahmed, A.; Bussey, H.; Fortin, N.; Friesen, J.D.; Storms, F.

submitted to the EMBL Data Library, December 1995

A:Description: The sequence of *Saccharomyces cerevisiae* chromosome XVI left arm.

A:Reference number: S69046

A:Accession: S69046

A:Molecule type: DNA

A:Residues: 1-460 <HAL>

A:CROSS-references: EMBL:U43703; NID:g1244769; PIDN:AA068221.1; PID:g1244776; MIPS:Y

C:Genetics:

A:Gene: SGD:UME1

A:CROSS-references: SGD:S0006060; MIPS:YPL139c

A:Map position: 16L

C:Superfamily: *Saccharomyces cerevisiae* transcription modulator WTM1

Query Match 64.3%; Score 36; DB 2; Length 460;  
 Best Local Similarity 62.5%; Pred. No. 25;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMHY 10  
 ||| |||  
 Db 85 IVPLGLHY 92

## RESULT 6

S38143

hypothetical protein YBL011w homolog YKR067w - yeast (*Saccharomyces cerevisiae*)

C:Species: *Saccharomyces cerevisiae*

C>Date: 03-May-1994 #sequence\_revision 03-May-1994 #text\_change 19-Apr-2002

C:Accession: S38143

R:van Vliet-Reedijk, J.C.; Planta, R.J.

submitted to the Protein Sequence Database, March 1994

A:Reference number: S38130

A:Accession: S38143

A:Molecule type: DNA

A:Residues: 1-743 <VAN>

A:CROSS-references: EMBL:Z28292; NID:g486536; PIDN:CAA82146.1; PID:g486537; MIPS:YKR

A:Experimental source: strain S288C

C:Genetics:

A:Gene: SGD:GPT2

A:CROSS-references: SGD:S0001775

A:Map position: 11R

C:Keywords: transmembrane protein

Query Match 64.3%; Score 36; DB 2; Length 743;  
 Best Local Similarity 75.0%; Pred. No. 43;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMHY 10  
||| |||  
Db 294 VVPCGLHY 301

RESULT 7  
D82618  
conserved hypothetical protein XFL1950 [imported] - Xylella fastidiosa (strain 9a5c)  
C:Species: Xylella fastidiosa  
C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 20-Aug-2000  
C:Accession: D82618  
R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequences  
Nature 406, 151-157, 2000  
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.  
A:Reference number: A82515; MUID:20365717; PMID:10910347  
A:Note: for a complete list of authors see reference number A59328 below  
A:Accession: D82618  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-156 <SIM>  
A:Cross-references: GB:AE004014; GB:AE003849; NID:99107044; PIDN:AAF84752.1; GSPDB:GN001  
A:Experimental source: strain 9a5c  
R:Simpon, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H. as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.  
submitted to Genbank, June 2000  
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E. A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.; F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A. Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveira M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z A:Reference number: A59328  
A:Contents: annotation  
C:Genetics:  
A:Gene: XFL1950

Query Match 62.5%; Score 35; DB 2; Length 156;  
Best Local Similarity 55.6%; Pred. No. 13;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
QY 1 EEVVPXGMH 9  
||| |||  
Db 119 EEILPQGVH 127

RESULT 8  
T02590  
DNA binding protein EREP-2 - common tobacco  
C:Species: Nicotiana tabacum (common tobacco)  
C:Date: 05-Mar-1999 #sequence\_revision 05-Mar-1999 #text\_change 21-Jul-2000  
C:Accession: T02590  
R:Ohme-Takagi, M.; Shinshi, H.  
Plant Cell 7, 173-182, 1995  
A:Title: Ethylene-inducible DNA binding proteins that interact with an ethylene response A:Reference number: Z14671; MUID:95276459; PMID:7756828  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-233 <OHM>  
A:Cross-references: EMBL:D38126; NID:9790362; PIDN:BAA07324.1; PID:gl208498  
A:Experimental source: strain BY4; tissue-type leaf

Query Match 62.5%; Score 35; DB 2; Length 233;  
Best Local Similarity 60.0%; Pred. No. 19;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
QY 1 EEVVPXGMHY 10  
: ||| |||  
Db 90 QAVVPKGRHY 99

RESULT 9  
H69194  
GMP synthetase, subunit B - Methanobacterium thermoautotrophicum (strain Delta H)  
C:Species: Methanobacterium thermoautotrophicum  
C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 22-Oct-1999  
C:Accession: H69194

R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jivan Ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N. J. Bacteriol. 179, 7135-7155, 1997  
A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: A:Reference number: A69000; MUID:98037514; PMID:9371463  
A:Accession: H69194  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-311 <MTH>  
A:Cross-references: GB:AE000850; GB:AE000666; NID:g2621794; PIDN:AA85215.1; PID:g2 A:Experimental source: strain Delta H  
C:Genetics:  
A:Gene: MTH710  
A:Start codon: GTG

Query Match 62.5%; Score 35; DB 2; Length 311;  
Best Local Similarity 63.6%; Pred. No. 27;  
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMHYS 11  
||||| |||  
Db 219 EEVVEGLHES 229

RESULT 10  
T24111  
hypothetical protein R10D12.10 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
C:Accession: T24111  
R:Percy, C.  
submitted to the EMBL Data Library, October 1996  
A:Reference number: Z19842  
A:Accession: T24111  
A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA  
A:Residues: 1-425 <WIL>  
A:Cross-references: EMBL:281109; PIDN:CAB03241.1; GSPDB:GN00023; CESP:R10D12.10 A:Experimental source: clone R10D12  
C:Genetics:  
A:Gene: CESP:R10D12.10  
A:Map position: 5  
A:Introns: 23/3; 56/3; 113/3; 257/2  
Query Match 62.5%; Score 35; DB 2; Length 425;  
Best Local Similarity 50.0%; Pred. No. 37;  
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMHY 10  
|:|:| |:  
Db 335 EQIVPGGLQY 344

RESULT 11

G86430  
T58.1 protein - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 17-May-2002  
C:Accession: G86430  
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; A. Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar ansen, N.F.; Hughes, B.; Huizar, L.  
Nature 408, 816-820, 2000  
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; K.

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziali, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
 A:Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.  
 A:Reference number: A86141; MUID:21016719; PMID:11130712  
 A:Accession: G86430  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-510 <STO>  
 A:Cross-references: GB:AE005172; NID:94587512; PIDN:AAD25743.1; GSPDB:GN00141  
 C:Genetics:  
 A:Map position: 1  
 C:Superfamily: hexose phosphate transport protein uhpT

Query Match 62.5%; Score 35; DB 2; Length 510;  
 Best Local Similarity 60.0%; Pred. No. 45;  
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMHY 10  
 |||||  
 Db 12 EEVKPPGIHF 21  
 |||||

RESULT 12  
 G69117  
 dipthine synthase - Methanobacterium thermoautotrophicum (strain Delta H)  
 C:Species: Methanobacterium thermoautotrophicum  
 C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 08-Oct-1999  
 C:Accession: G69117  
 R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.; Olu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiواني, N. ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.  
 J. Bacteriol. 179, 7135-7155, 1997  
 A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funct  
 A:Reference number: A69000; MUID:98037514; PMID:9371463  
 A:Accession: G69117  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-264 <TH>  
 A:Cross-references: GB:AE000940; GB:AE000666; NID:92623011; PIDN:AAB86340.1; PID:G262301  
 A:Experimental source: strain Delta H  
 C:Genetics:  
 A:Gene: MTH1874  
 C:Superfamily: Methanococcus jannaschii dipthine synthase

Query Match 60.7%; Score 34; DB 2; Length 264;  
 Best Local Similarity 62.5%; Pred. No. 35;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMHY 10  
 |||||  
 Db 235 VVPAGLHF 242  
 |||||

RESULT 13  
 C75538  
 hypothetical protein - Deinococcus radiodurans (strain R1)  
 C:Species: Deinococcus radiodurans  
 C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 28-Jul-2000  
 C:Accession: C75538  
 R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.  
 Science 286, 1571-1577, 1999  
 A:Title: Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1.  
 A:Reference number: A75250; MUID:20036896; PMID:10567266  
 A:Accession: C75538  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-279 <WHI>  
 A:Cross-references: GB:AE001889; GB:AE000513; NID:G6457944; PIDN:AAF09867.1; PID:G645795  
 A:Experimental source: strain R1

## C:Genetics:

A:Gene: DR0271

A:Map position: 1

C:Superfamily: Deinococcus radiodurans hypothetical protein DR0271

Query Match 60.7%; Score 34; DB 2; Length 279;

Best Local Similarity 75.0%; Pred. No. 38;

Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 VVPXGMHY 11  
|||Db 100 VPLGRHYS 107  
|||

## RESULT 14

B75478

3-dehydroquininate synthase - Deinococcus radiodurans (strain R1)

C:Species: Deinococcus radiodurans

C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 17-Mar-2000

C:Accession: B75478

R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.

M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.

S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.

Science 286, 1571-1577, 1999

A:Title: Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1.

A:Reference number: A75250; MUID:20036896; PMID:10567266

A:Accession: B75478

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-350 &lt;WHI&gt;

A:Cross-references: GB:AE001932; GB:AE000513; NID:G6458481; PIDN:AAF10353.1; PID:G64

A:Experimental source: strain R1

C:Genetics:

A:Gene: DR0777

A:Map position: 1

C:Superfamily: 3-dehydroquininate synthase; 3-dehydroquininate synthase homology

Query Match 60.7%; Score 34; DB 2; Length 350;

Best Local Similarity 60.0%; Pred. No. 48;

Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 EVVPXGMHY 11  
|||Db 252 EAVYGMHYA 261  
|||

## RESULT 15

T35025

probable DNA ligase - Streptomyces coelicolor

C:Species: Streptomyces coelicolor

C:Date: 05-Nov-1999 #sequence\_revision 05-Nov-1999 #text\_change 05-Nov-1999

C:Accession: T35025

R:Seeger, S.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.

submitted to the EMBL Data Library, June 1999

A:Reference number: 221565

A:Accession: T35025

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-355 &lt;SEE&gt;

A:Cross-references: EMBL:AL079355; PIDN:CAB45581.1; GSPDB:GN00070; SCOEDB:SC4C6.17C

A:Experimental source: strain A3(2)

C:Genetics:

A:Gene: SCOEDB:SC4C6.17C

Query Match 60.7%; Score 34; DB 2; Length 355;

Best Local Similarity 71.4%; Pred. No. 49;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 VVPXGMHY 10  
|||Db 20 IPPGMHY 26  
|||

A:Cross-references: EMBL:Z50154; NID:gl052827; PID:gl052828

Query Match 60.7%; Score 34; DB 2; Length 426;  
Best Local Similarity 44.4%; Pred. No. 59;  
Matches 4; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXGMH 9

Db 52 DQVIPAGLH 60

RESULT 19

T28717

hypothetical protein F10D2.3 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 29-Oct-1999

C:Accession: T28717

R:Graves, T.; Wohldmann, P.; Gillam, B.

A:Description: The sequence of C. elegans cosmid F10D2.

A:Reference number: Z20515

A:Accession: T28717

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-495 <GRA>

A:Cross-references: EMBL:AF022972; PIDN:AAC48234.1; GSPDB:GN000023; CESP:F10D2.3

A:Experimental source: strain Bristol N2; clone F10D2

C:Genetics:

A:Gene: CESP:F10D2.3

A:Map position: 5

A:Introns: 37/2; 90/1; 113/1; 183/3; 356/3; 380/3; 428/2

Query Match 60.7%; Score 34; DB 2; Length 495;

Best Local Similarity 50.0%; Pred. No. 70;

Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMHY 10

Db 218 ENIVPTGKH 227

RESULT 20

AF3286

ATP-dependent DNA helicase BMEI0275 [imported] - Brucella melitensis (strain 16M)

C:Species: Brucella melitensis

C>Date: 01-Feb-2002 #sequence\_revision 01-Feb-2002 #text\_change 01-Feb-2002

C:Accession: AF3286

R:DelVecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivar

Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002

A:Title: The genome sequence of the facultative intracellular pathogen Brucella mel

A:Reference number: AD3252; PMID:11756688

A:Accession: AF3286

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1028 <KUR>

A:Cross-references: GB:AE008917; PIDN:AAL51457.1; PID:gl17982167; GSPDB:GN00190

A:Experimental source: strain 16M

C:Genetics:

A:Gene: BMEI0275

A:Map position: 1

Query Match 60.7%; Score 34; DB 2; Length 1028;

Best Local Similarity 54.5%; Pred. No. 1.5e+02;

Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMHYS 11

Db 76 EKIVPPGARYS 86

RESULT 21

S54619

RESULT 16

E69086

cell division protein - Methanobacterium thermoautotrophicum (strain Delta H)

C:Species: Methanobacterium thermoautotrophicum

C>Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 29-Sep-1999

C:Accession: E69086

R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;

Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwan, N.

kl, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.

J. Bacteriol. 179, 7135-7155, 1997

A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funct

A:Reference number: A69000; MUID:98037514; PMID:9371463

A:Accession: E69086

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-360 <MTH>

A:Cross-references: GB:AE000923; GB:AE000666; NID:g2622766; PIDN:AAB86115.1; PID:g262277

A:Experimental source: strain Delta H

C:Genetics:

A:Gene: MTH1642

C:Superfamily: cell division protein MJ0174

Query Match 60.7%; Score 34; DB 2; Length 360;

Best Local Similarity 45.5%; Pred. No. 50;

Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMHYS 11

Db 98 EDLVPNGSHRT 108

RESULT 17

C83903

hypothetical protein BH2027 [imported] - Bacillus halodurans (strain C-125)

C:Species: Bacillus halodurans

C>Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 15-Jun-2001

C:Accession: C83903

R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira

Nucleic Acids Res. 28, 4317-4331, 2000

A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and

A:Reference number: A83650; MUID:20512582; PMID:11058132

A:Accession: C83903

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-425 <STO>

A:Cross-references: GB:AP001514; GB:BA000004; NID:gl0174613; PIDN:BA805746.1; GSPDB:GN00

A:Experimental source: strain C-125

C:Genetics:

A:Gene: BH2027

Query Match 60.7%; Score 34; DB 2; Length 425;

Best Local Similarity 60.0%; Pred. No. 59;

Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMHY 10

Db 403 EELLIEGMHY 412

RESULT 18

S58132

Sls1 protein precursor - yeast (Yarrowia lipolytica)

C:Species: Yarrowia lipolytica, Candida lipolytica

C>Date: 13-Jan-1996 #sequence\_revision 01-Mar-1996 #text\_change 09-Sep-1997

C:Accession: S58132

R:Boisrame, A.; Beckerich, J.; Gaillardin, C.

submitted to the EMBL Data Library, July 1995

A:Description: Sls1p, an endoplasmic reticulum component, is involved in the protein tra

A:Reference number: S58132

A:Accession: S58132

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-426 <BOI>

hypothetical protein YOR013w - yeast (*Saccharomyces cerevisiae*)  
 N/Alternate names: hypothetical protein O2612; hypothetical protein YOL303.3  
 C/Species: *Saccharomyces cerevisiae*  
 C/Date: 08-Jul-1995 #sequence\_revision 01-Sep-1995 #text\_change 19-Apr-2002  
 C/Accession: S54619; S66879  
 R/de Haan, M.; Maarse, A.C.; Grivell, L.A.  
 submitted to the EMBL Data Library, May 1995  
 A/Reference number: S54617  
 A/Accession: S54619  
 A/Molecule type: DNA  
 A/Residues: 1-156 <DEH>  
 A/Cross-references: EMBL:X87331; NID:g1041652; PIDN:CAA60762.1; PID:g829123  
 R/de Haan, M.; Grivell, L.A.; Maarse, A.C.  
 submitted to the Protein Sequence Database, July 1996  
 A/Reference number: S66877  
 A/Accession: S66879  
 A/Molecule type: DNA  
 A/Residues: 1-156 <DEW>  
 A/Cross-references: EMBL:Z74920; NID:g1420109; PIDN:CAA99201.1; PID:g1420111; MIPS:YOR013w  
 A/Experimental source: strain S288C  
 C/Genetics:  
 A/Cross-references: SGD:S0005539  
 A/Map position: 15R  
 C/Superfamily: hypothetical protein YOR013w

Query Match 58.9%; Score 33; DB 2; Length 156;  
 Best Local Similarity 66.7%; Pred. No. 32;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 EVVVPXGMHY 10  
 ||| |||  
 Db 50 EVVPLGNDY 58

# RESULT 22

E83607  
 polyamine transport protein PA0300 [imported] - *Pseudomonas aeruginosa* (strain PA01)  
 C/Species: *Pseudomonas aeruginosa*  
 C/Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
 C/Accession: E83607  
 R/Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Bradman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, J.; Lory, S.; Olson, M.V.  
 Nature 406, 959-964, 2000  
 A/Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pathogen  
 A/Reference number: A82950; MUID:20437337; PMID:10984043  
 A/Accession: E83607  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-367 <STO>  
 A/Cross-references: GB:AE004468; GB:AE004091; NID:g9946144; PIDN:AAG03689.1; GSPDB:GN001  
 A/Experimental source: strain PA01  
 C/Genetics:  
 A/Gene: potF2; PA0300  
 C/Superfamily: *Escherichia coli* spermidine/putrescine-binding protein

Query Match 58.9%; Score 33; DB 2; Length 367;  
 Best Local Similarity 44.4%; Pred. No. 81;  
 Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 EVVVPXGMHY 10  
 ||| |||  
 Db 183 EILPAALHY 191

# RESULT 23

G82253  
 conserved hypothetical protein VC1005 [imported] - *Vibrio cholerae* (strain N16961 serotype O1)  
 C/Species: *Vibrio cholerae*  
 C/Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Feb-2001  
 C/Accession: G82253  
 R/Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.; Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, B.

l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
 Nature 406, 477-483, 2000  
 A/Title: DNA sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.  
 A/Reference number: A82035; MUID:20406833; PMID:10952301  
 A/Accession: G82253  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-441 <HEI>  
 A/Cross-references: GB:AE004182; GB:AE003852; NID:g9655461; PIDN:AAF94166.1; GSPDB:G000000000  
 A/Experimental source: serogroup O1; strain N16961; biotype El Tor  
 C/Genetics:  
 A/Gene: VC1005  
 A/Map position: 1  
 C/Superfamily: conserved integral membrane protein HP0758

Query Match 58.9%; Score 33; DB 2; Length 441;  
 Best Local Similarity 54.5%; Pred. No. 98;  
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 EVVVPXGMHY 11  
 ||| |||  
 Db 417 ETVVPTFIHYN 427

# RESULT 24

G71542  
 probable amino acid transporter - *Chlamydia trachomatis* (serotype D, strain UW3/Cx)  
 C/Species: *Chlamydia trachomatis*  
 C/Date: 13-Sep-1998 #sequence\_revision 13-Sep-1998 #text\_change 19-May-2000  
 C/Accession: G71542  
 R/Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchell, D.; White, O.; Nelson, M.; Nelson, W.; deBoy, R.; Kolonay, J.; McClarty, G.; Salzberg, S.L.; Fraser, C.M.; Venter, J.C.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
 Science 282, 754-759, 1998  
 A/Title: Genome sequence of an obligate intracellular pathogen of humans: *Chlamydia trachomatis*  
 A/Reference number: A71570; MUID:99000809; PMID:9784136  
 A/Accession: G71542  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-466 <ARN>  
 A/Cross-references: GB:AE001295; GB:AE001273; NID:g3328617; PIDN:AAC67808.1; PID:g3328617  
 A/Experimental source: serotype D, strain UW-3/Cx  
 C/Genetics:  
 A/Gene: xsa  
 C/Superfamily: L-lysine transport protein

Query Match 58.9%; Score 33; DB 2; Length 466;  
 Best Local Similarity 62.5%; Pred. No. 1e+02;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 4 VPXGMHY 11  
 ||| |||  
 Db 453 IPFGMYYS 460

# RESULT 25

H81697  
 amino acid antiporter TC0488 [imported] - *Chlamydia muridarum* (strain N19g)  
 C/Species: *Chlamydia muridarum*, *Chlamydia trachomatis* MoPn  
 C/Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 02-Jun-2000  
 C/Accession: H81697  
 R/Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey, M.J.; Nelson, M.; Nelson, W.; deBoy, R.; Kolonay, J.; McClarty, G.; Salzberg, S.L.; Fraser, C.M.; Venter, J.C.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
 Nucleic Acids Res. 28, 1397-1406, 2000  
 A/Title: Genome sequences of *Chlamydia trachomatis* MoPn and *Chlamydia pneumoniae* AR39  
 A/Reference number: A81500; MUID:20150255; PMID:10684935  
 A/Accession: H81697  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-466 <TET>  
 A/Cross-references: GB:AE002317; GB:AE002160; NID:g7190522; PIDN:AAF39334.1; PID:g7190522  
 A/Experimental source: strain N19g (MoPn)  
 C/Genetics:  
 A/Gene: TC0488  
 C/Superfamily: L-lysine transport protein



us-09-909-164-11.rpr

Wed Jun 11 15:42:19 2003

Query Match 58.9%; Score 33; DB 2; Length 466;  
 Best Local Similarity 62.5%; Pred. No. 1e+02;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 4 VPXGMHYS 11  
 : | | | | |  
 Db 453 IPFGMYYS 460

Search completed: June 10, 2003, 13:49:14  
 Job time : 12.2143 secs



GenCore version 5.1.6  
Copyright (C) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2003, 13:25:04 ; Search time 4.5 Seconds  
(without alignments)  
101.387 Million cell updates/sec

Title: US-09-909-164-11  
Perfect score: 56  
Sequence: 1 EREVPGMHYS 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 75 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	38	67.9	1058	1	Q8986 fusbacteri
2	37	66.1	1057	1	Q99ur5 staphylococ
3	37	66.1	1057	1	P58940 staphylococ
4	36	64.3	102	1	P31619 tobacco yel
5	36	64.3	460	1	Q0310 saccharomyc
6	36	64.3	743	1	P36148 saccharomyc
7	35	62.5	227	1	Q35586 mesocricetu
8	35	62.5	308	1	O26806 methanobact
9	34	60.7	426	1	Q99158 yarrowia li
10	33	58.9	513	1	P13065 desulfovibr
11	33	58.9	627	1	P49850 bacillus su
12	33	58.9	1188	1	P34335 caenorhabdi
13	33	58.9	1396	1	P13390 bacterioph
14	32	57.1	126	1	P08858 bos taurus
15	32	57.1	267	1	P56351 chlorella v
16	32	57.1	288	1	Q04827 rattus norv
17	32	57.1	289	1	P30279 homo sapien
18	32	57.1	289	1	P30280 mus musculu
19	32	57.1	291	1	P50755 xenopus lae
20	32	57.1	291	1	P49706 gallus gall
21	32	57.1	291	1	P53782 xenopus lae
22	32	57.1	291	1	P55169 gallus gall
23	32	57.1	292	1	P30281 homo sapien
24	32	57.1	292	1	P24385 homo sapien
25	32	57.1	295	1	P25322 mus musculu
26	32	57.1	295	1	P25322 mus musculu
27	32	57.1	295	1	P40595 azotobacter
28	32	57.1	341	1	P19887 bacillus an
29	32	57.1	353	1	P32784 saccharomyc
30	32	57.1	759	1	P13784 saccharomyc
31	32	57.1	877	1	Q74377 schizosacch
32	32	57.1	1401	1	Q9Kv29 vibrio chol
33	32	57.1	2717	1	P15822 homo sapien

34	31.5	56.2	847	1	CD222_HUMAN
35	31	55.4	124	1	REV_SIVCZ
36	31	55.4	130	1	SZ05_RAT
37	31	55.4	276	1	Y939_METJA
38	31	55.4	319	1	YHAI_CRYPA
39	31	55.4	331	1	RL3_ARCFU
40	31	55.4	363	1	Y325_HAEIN
41	31	55.4	450	1	ENP3_HUMAN
42	31	55.4	529	1	PTBA_BACSU
43	31	55.4	609	1	DCMB_MOOTH
44	31	55.4	674	1	ETRI_BRAOL
45	31	55.4	735	1	ETRI_ARATH
46	31	55.4	738	1	CY14_NEUCR
47	31	55.4	788	1	CENC_MOUSE
48	31	55.4	906	1	Q09541 caenorhabdi
49	31	55.4	1374	1	Q09541 caenorhabdi
50	31	55.4	1378	1	Q09541 caenorhabdi
51	31	55.4	1394	1	Q09541 caenorhabdi
52	31	55.4	1429	1	Q09541 caenorhabdi
53	31	55.4	1498	1	Q09541 caenorhabdi
54	31	55.4	1595	1	Q09541 caenorhabdi
55	31	55.4	1712	1	Q09541 caenorhabdi
56	30.5	54.5	73	1	IF1_CHLPN
57	30.5	54.5	73	1	IF1_CHLTPR
58	30.5	54.5	492	1	CATL_ARATH
59	30	53.6	121	1	NU4M_TRIRU
60	30	53.6	121	1	TKNK_HUMAN
61	30	53.6	152	1	Y16D_BPT4
62	30	53.6	172	1	YKRB_BACSU
63	30	53.6	185	1	PAGC_SALTY
64	30	53.6	212	1	MSRA_VIBCH
65	30	53.6	219	1	SSL_DROME
66	30	53.6	223	1	COAT_CTV36
67	30	53.6	232	1	SCOA_HELPJ
68	30	53.6	232	1	SCOA_HELPY
69	30	53.6	298	1	YA52_HAEIN
70	30	53.6	331	1	LDHC_HUMAN
71	30	53.6	358	1	YVAA_BACSU
72	30	53.6	363	1	PITL_CHICK
73	30	53.6	370	1	PITL_MELGA
74	30	53.6	396	1	DHL_HUMAN
75	30	53.6	396	1	DHL_MOUSE

## ALIGNMENTS

## RESULT 1

ID	CARB_FUSNN	STANDARD	PRT	1058 AA
AC	Q8RG86			
DT	15-JUN-2002 (Rel. 41, Created)			
DT	15-JUN-2002 (Rel. 41, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Carbamoyl-phosphate synthase large chain (EC 6.3.5.5) (Carbamoyl-phosphate synthase ammonia chain).			
DE	CARB OR FN0422.			
GN	Fusobacterium nucleatum (subsp. nucleatum).			
OS	Fusobacterium nucleatum (subsp. nucleatum).			
OC	Bacteria; Fusobacteria; Fusobacterium.			
OX	NCBI_TaxID=76856;			
RN	[1]			
RC	SEQUENCE FROM N.A.			
RC	STRAIN=ATCC 25586;			
RX	MEDLINE=21886394; PubMed=11889109;			
RA	Kapatal V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A.,			
RA	Bhattacharyya A., Bartman A., Gardner W., Grechkin G., Zhu L.,			
RA	Vasileva O., Chu L., Kogan Y., Chaga O., Goldsman E., Bernal A.,			
RA	Larsen N., D'Souza M., Walunas T., Fusch G., Haselkorn R.,			
RA	Fonstein M., Kyrides N., Overbeek R.,			
RT	"Genome sequence and analysis of the oral bacterium Fusobacterium			
RT	nucleatum strain ATCC 25586."			
RL	Nucleotide 184:2005-2018(2002).			
CC	-I- CATALYTIC ACTIVITY: 2 ATP + L-glutamine + CO(2) + H(2)O = 2 ADP +			

CC phosphate + L-glutamate + carbamoyl phosphate.  
 CC -1- COFACTOR: Arginine biosynthesis.  
 CC -1- PATHWAY: Pyrimidine biosynthesis.  
 CC -1- SUBUNIT: Composed of two chains; the small (or glutamine) chain  
 CC promotes the hydrolysis of glutamine to ammonia, which is used by  
 CC the large (or ammonia) chain to synthesize carbamoyl phosphate (By  
 CC similarity).  
 CC -1- SIMILARITY: BELONGS TO THE CARB FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: AE010554; AAL94625.1; ALT\_INIT.  
 CC InterPro: IPR005483; CPase\_L.  
 CC InterPro: IPR005479; CPase\_L\_D2.  
 CC InterPro: IPR005480; CPase\_L\_D3.  
 CC InterPro: IPR005481; CPase\_L\_N.  
 CC InterPro: IPR004362; MGS-like.  
 CC Pfam: PF00289; CPase\_L\_chain; 2.  
 CC Pfam: PF02786; CPase\_L\_D2; 2.  
 CC Pfam: PF02787; CPase\_L\_D3; 1.  
 CC Pfam: PF02142; MGS; 1.  
 CC PRINTS: PR00098; CPASE.  
 CC PROSITE: PS00866; CPASE\_1; 2.  
 CC PROSITE: PS00867; CPASE\_2; 2.  
 CC Arginine biosynthesis; Pyrimidine biosynthesis; Ligase; Repeat;  
 KW ATP-binding; Manganese; Complete proteome.  
 FT DOMAIN 1 401 CARBOXYPHOSPHATE SYNTHETIC DOMAIN.  
 FT DOMAIN 402 546 OLIGOMERIZATION DOMAIN.  
 FT DOMAIN 547 929 CARBAMOYL PHOSPHATE SYNTHETIC DOMAIN.  
 FT REPEAT 1 546 ALLOSTERIC DOMAIN.  
 FT REPEAT 547 1058  
 FT NP\_BIND 153 210 ATP (POTENTIAL).  
 FT NP\_BIND 302 352 ATP (POTENTIAL).  
 FT METAL 284 298 MANGANESE 1 (BY SIMILARITY).  
 FT METAL 298 298 MANGANESE 1 AND 2 (BY SIMILARITY).  
 FT METAL 300 300 MANGANESE 2 (BY SIMILARITY).  
 FT METAL 820 820 MANGANESE 3 (BY SIMILARITY).  
 FT METAL 832 832 MANGANESE 3 (BY SIMILARITY).  
 SQ SEQUENCE 1058 AA; 117451 MW; ED7037AF77C1E39F CRC64;  
 Query Match 67.9%; Score 38; DB 1; Length 1058;  
 Best Local Similarity 60.0%; Pred. No. 10;  
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
 QY 2 EVVPGMHYS 11  
 Db 190 EIVPGLNYS 199  
 RESULT 2  
 CARB\_STAAM STANDARD; PRT; 1057 AA.  
 AC Q99UR5;  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Carbamoyl-phosphate synthase large chain (EC 6.3.5.5) (Carbamoyl-  
 DE phosphate synthetase ammonia chain).  
 GN CARB OR PYRAB OR SAV1203 OR SAI046.  
 OS Staphylococcus aureus (strain N315) / ATCC 700699), and  
 OS Bacteria; Firmicutes; Bacillales; Staphylococcus.  
 OC NCBI\_TaxID=158878, 158879;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RC STRAIN-Mu50 / ATCC 700699, and N315;  
 RX MEDLINE=213111952; PubMed=11418146;  
 RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,  
 RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,  
 RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,  
 RA Mizutani-U Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,  
 RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,  
 RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,  
 RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;  
 FT "Whole genome sequencing of methicillin-resistant Staphylococcus  
 FT aureus".  
 RL Lancet 357:1325-1240(2001).  
 CC -1- CATALYTIC ACTIVITY: 2 ATP + L-glutamine + CO(2) + H(2)O = 2 ADP +  
 CC phosphate + L-glutamate + carbamoyl phosphate.  
 CC -1- COFACTOR: Binds three manganese ions (By similarity).  
 CC -1- PATHWAY: Arginine biosynthesis.  
 CC -1- PATHWAY: Pyrimidine biosynthesis; first step.  
 CC -1- SUBUNIT: Composed of two chains; the small (or glutamine) chain  
 CC promotes the hydrolysis of glutamine to ammonia, which is used by  
 CC the large (or ammonia) chain to synthesize carbamoyl phosphate (By  
 CC similarity).  
 CC -1- SIMILARITY: BELONGS TO THE CARB FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: AP003361; BAB57365.1; -  
 CC EMBL: AP003132; BAB42298.1; -  
 CC HSPSP; P00968; 1CS0.  
 CC InterPro: IPR005483; CPase\_L.  
 CC InterPro: IPR005479; CPase\_L\_D2.  
 CC InterPro: IPR005480; CPase\_L\_D3.  
 CC InterPro: IPR005481; CPase\_L\_N.  
 CC InterPro: IPR004362; MGS-like.  
 CC Pfam: PF00289; CPase\_L\_chain; 2.  
 CC Pfam: PF02786; CPase\_L\_D2; 2.  
 CC Pfam: PF02787; CPase\_L\_D3; 1.  
 CC PRINTS: PR00098; CPASE.  
 CC PROSITE: PS00866; CPASE\_1; 2.  
 CC PROSITE: PS00867; CPASE\_2; 2.  
 CC Arginine biosynthesis; Pyrimidine biosynthesis; Ligase; Repeat;  
 KW ATP-binding; Manganese; Complete proteome.  
 FT DOMAIN 1 401 CARBOXYPHOSPHATE SYNTHETIC DOMAIN.  
 FT DOMAIN 402 546 OLIGOMERIZATION DOMAIN.  
 FT DOMAIN 547 929 CARBAMOYL PHOSPHATE SYNTHETIC DOMAIN.  
 FT REPEAT 1 546 ALLOSTERIC DOMAIN.  
 FT REPEAT 547 1057  
 FT NP\_BIND 153 210 ATP (POTENTIAL).  
 FT NP\_BIND 302 352 ATP (POTENTIAL).  
 FT METAL 284 298 MANGANESE 1 (BY SIMILARITY).  
 FT METAL 298 298 MANGANESE 1 AND 2 (BY SIMILARITY).  
 FT METAL 300 300 MANGANESE 2 (BY SIMILARITY).  
 FT METAL 820 820 MANGANESE 3 (BY SIMILARITY).  
 FT METAL 832 832 MANGANESE 3 (BY SIMILARITY).  
 SQ SEQUENCE 1057 AA; 117171 MW; E3E179EF0591F0F8 CRC64;  
 Query Match 66.1%; Score 37; DB 1; Length 1057;  
 Best Local Similarity 60.0%; Pred. No. 16;  
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
 QY 2 EVVPGMHYS 11  
 Db 190 EIVSNGHYS 199  
 RESULT 3

```

CARB_STRAW
ID CARB_STRAW STANDARD; PRT; 1057 AA.
AC P58940;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Carbamoyl-phosphate synthase large chain (EC 6.3.5.5) (Crbamoyl-
GN CARB OR PYRAB OR MW1086.
OS Staphylococcus aureus (strain MW2).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=196620;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22040717; PubMed=12044378;
RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
RA Nagai Y., Iwama N., Asano K., Naimi T., Kuroda H., Cui L.,
RA Yamamoto K., Hiramatsu K.;
RT "Genome and virulence determinants of high virulence community-
acquired MRSA.";
RL Lancet 359:1819-1827(2002).
CC -|- CATALYTIC ACTIVITY: 2 ATP + L-glutamine + CO(2) + H(2)O = 2 ADP +
phosphate + L-glutamate + carbamoyl phosphate.
CC -|- COFACTOR: Binds three manganese ions (By similarity).
CC -|- PATHWAY: Arginine biosynthesis; first step.
CC -|- PATHWAY: Pyrimidine biosynthesis; the small (or glutamine) chain
promotes the hydrolysis of glutamine to ammonia, which is used by
the large (or ammonia) chain to synthesize carbamoyl phosphate (By
similarity).
CC -|- SIMILARITY: BELONGS TO THE CARB FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announcement/
or send an email to license@isb-sib.ch).
CC
CC EMBL; AF004825; BAB94951.1; -.
DR PROSITE; PS00866; CPSASE_1; 2.
DR PROSITE; PS00867; CPSASE_2; 2.
KW Arginine biosynthesis; Pyrimidine biosynthesis; Ligase; Repeat;
KW ATP-binding; Manganese.
FT DOMAIN 1 401 CARBOXYPHOSPHATE SYNTHETIC DOMAIN.
FT DOMAIN 402 546 OLIGOMERIZATION DOMAIN.
FT DOMAIN 547 929 CARBAMOYL PHOSPHATE SYNTHETIC DOMAIN.
FT DOMAIN 930 1057 ALLOSTERIC DOMAIN.
FT REPEAT 1 546
FT REPEAT 547 1057
FT NP_BIND 153 210 ATP (POTENTIAL).
FT NP_BIND 302 352 ATP (POTENTIAL).
FT METAL 284 284 MANGANESE 1 (BY SIMILARITY).
FT METAL 298 298 MANGANESE 1 AND 2 (BY SIMILARITY).
FT METAL 300 300 MANGANESE 2 (BY SIMILARITY).
FT METAL 820 820 MANGANESE 3 (BY SIMILARITY).
FT METAL 832 832 MANGANESE 3 (BY SIMILARITY).
SQ SEQUENCE 1057 AA; 117185 MW; D8E3B09F9BC6F152 CRC64;

Query Match 66.1%; Score 37; DB 1; Length 1057;
Best Local Similarity 60.0%; Pred. No. 16;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 EVVPGMHYS 11
|:| |::|
Db 190 EIVSNGLHYS 199

RESULT 4
Y1LK_TYDVA STANDARD; PRT; 102 AA.
ID Y1LK_TYDVA
AC P31619;

```

```

DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 01-OCT-1993 (Rel. 27, Last annotation update)
DE Hypothetical 11.2 kDa protein.
GN V1.
OS Tobacco yellow dwarf virus (strain Australia) (TYDV).
OC Viruses; ssDNA viruses; Geminiviridae; Mastrevirus.
OX NCBI_TaxID=31599;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92188538; PubMed=1546458;
RA Morris B.A.M., Richardson K.A., Haley A., Zhan X., Thomas J.E.;
RT "The nucleotide sequence of the infectious cloned DNA component of
to tobacco yellow dwarf virus reveals features of geminiviruses
inflicting monocotyledonous plants.";
RL Virology 187:633-642(1992).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announcement/
or send an email to license@isb-sib.ch).
CC
CC EMBL; M81103; AAA47947.1; -.
DR PIR; A42452; A42452.
DR InterPro; IPR002621; Gemini_mov.
DR Pfam; PF01708; Gemini_mov; 1.
KW Hypothetical protein.
KW
SQ SEQUENCE 102 AA; 111178 MW; A40ECF1E0AF5B67 CRC64;

Query Match 64.3%; Score 36; DB 1; Length 102;
Best Local Similarity 60.0%; Pred. No. 2.3;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 EVVPGMHYS 11
|:| |::|
Db 7 QVVPSSGINS 16

RESULT 5
UMEL_YEAST STANDARD; PRT; 450 AA.
ID UMEL_YEAST
AC Q03010; P87330;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Meiosis negative regulator UMEL.
GN UMEL OR WTM3 OR YPL139C OR LPI7C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A364A;
RA Mallory M.J., Strich R.;
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c / AB972;
RX MEDLINE=97313271; PubMed=9169875;
RA Bussey H., Storms R.K., Ahmed A., Albermann K., Allen E., Ansorge W.,
RA Araujo R., Aparicio A., Barrell B.G., Badcock K., Benes V.,
RA Botstein D., Bowman S., Bruckner M., Carpenter J., Cherry J.M.,
RA Chung E., Churcher C.M., Coster F., Davis K., Davis R.W.,
RA Dietrich F.S., Dillus H., Difaolo T., Dubois E., Duesterhoeft A.,
RA Duncan M., Floeth U., Fortin N., Friesen J.D., Fritz C., Goffeau A.,
RA Hall J., Hebling U., Heumann K., Hilbert H., Hillier L.,
RA Hunicker-Smith S., Hyman R., Johnston M., Kalman S., Kleine K.,
RA Komp C., Kurdi O., Lashkari D., Lew H., Lin A., Lin D., Louis E.J.,
RA Marathe R., Messenguy F., Mewes H.-W., Mirtipati S., Moestl D.,

```

RA Mueller-Auer S., Namath A., Kentwich U., Oefner P., Pearson D.,  
 RA Petel F.X., Fohl T.M., Purnelle D., Schafer M., Scharfe M.,  
 RA Scherens B., Schramm S., Schroeder M., Sdicu A.M., Tettelin H.,  
 RA Urrestarazu L.A., Ushinsky S., Vierendeels F., Vissers S., Voss H.,  
 RA Walsh S.V., Wambutt R., Wang Y., Wedler E., Wedler H., Winnett E.,  
 RA Zhong W.W., Zollner A., Vo D.H., Hani J.,  
 RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome XVI."  
 RL Nature 387:103-105(1997).  
 CC -1- FUNCTION: TRANSCRIPTIONAL MODULATOR WITH ROLES IN MEIOTIC  
 CC REGULATION AND SILENCING. NEGATIVE REGULATOR OF MEIOSIS.  
 CC -1- SIMILARITY: CONTAINS 4 WD REPEATS (TRP-ASP DOMAINS).  
 CC -1- SIMILARITY: STRONG, TO YEAST WTML AND WTM2.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; U10280; AAB40937.1; -;  
 DR EMBL; U43703; AAB68221.1; -;  
 DR TRANSFAC; T04309; -;  
 DR SGD; S0006060; UME1.  
 DR InterPro; IPR001680; WD40.  
 DR Pfam; PF00400; WD40; 3.  
 DR SMART; SM00320; WD40; 3.  
 DR PROSITE; PS00678; WD\_REPEATS.1; FALSE\_NEG.  
 DR PROSITE; PS00882; WD\_REPEATS.2; FALSE\_NEG.  
 DR PROSITE; PS0294; WD\_REPEATS\_REGION; FALSE\_NEG.  
 KW Transcription regulation; Meiosis; Repeat; WD repeat.  
 FT REPEAT 233 271 WD 1.  
 FT REPEAT 276 316 WD 2.  
 FT REPEAT 339 379 WD 3.  
 FT REPEAT 411 451 WD 4.  
 FT REPEAT 460 510 WD 4.  
 SQ SEQUENCE 460 AA; 51022 MW; AA6F60448B7BCBA9 CRC64;  
 Query Match 64.3%; Score 36; DB 1; Length 460;  
 Best Local Similarity 62.5%; Pred. No. 11;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 QY 3 VVPXGMHY 10  
 Db 85 IVPGLGHY 92  
 RESULT 6  
 YK47\_YEAST STANDARD; PRT; 743 AA.  
 AC P36148;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 01-NOV-1997 (Rel. 35, Last annotation update)  
 DE Hypothetical 83.6 kDa protein in CCP1-MET1 intergenic region.  
 GN YK067W.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
 OX NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=S288c;  
 RA van Vliet-Reedijk J.C., Planta R.J.;  
 RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Probable).  
 CC -1- SIMILARITY: STRONG, TO YEAST YBL011W.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; Z28292; CAA82146.1; -;  
 DR PIR; S38143; S38143.  
 DR SGD; S0001775; YK067W.  
 DR InterPro; IPR002123; Acyltransferase.  
 KW Hypothetical protein; Transmembrane.  
 FT TRANSMEM 31 55 POTENTIAL.  
 FT TRANSMEM 69 85 POTENTIAL.  
 FT TRANSMEM 502 524 POTENTIAL.  
 FT TRANSMEM 539 555 POTENTIAL.  
 SQ SEQUENCE 743 AA; 83644 MW; 84B9946E56B82F15 CRC64;  
 Query Match 64.3%; Score 36; DB 1; Length 743;  
 Best Local Similarity 75.0%; Pred. No. 18;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 3 VVPXGMHY 10  
 Db 294 VVPCGLHY 301  
 RESULT 7  
 IDIL\_MESAU STANDARD; PRT; 227 AA.  
 AC O35586;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Isopentenyl-diphosphate delta-isomerase 1 (EC 5.3.3.2) (IPP isomerase  
 DE 1) (Isopentenyl pyrophosphate isomerase 1) (IPPI).  
 GN IDIL.  
 OS Mesocricetus auratus (Golden hamster).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;  
 OC Mesocricetus.  
 OX NCBI\_TaxID=10036;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97373600; PubMed=9228075;  
 RA Paton V.G., Shackelford J.E., Krisans S.K.;  
 RT Cloning and subcellular localization of hamster and rat isopentenyl  
 RT diphosphate dimethylallyl diphosphate isomerase. A PTS1 motif targets  
 RT the enzyme to peroxisomes.\*;  
 RL J. Biol. Chem. 272:18945-18950(1997).  
 CC -1- FUNCTION: CATALYZES THE 1,3-ALLYLIC REARRANGEMENT OF THE  
 CC HOMOALLYLIC SUBSTRATE ISOPENTENYL (IPP) TO ITS HIGHLY  
 CC ELECTROPHILIC ALLYLIC ISOMER, DIMETHYLLALLYL DIPHOSPHATE (DMAPP).  
 CC -1- CATALYTIC ACTIVITY: Isopentenyl diphosphate = dimethylallyl  
 CC diphosphate.  
 CC -1- COFACTOR: REQUIRES MAGNESIUM FOR ACTIVITY.  
 CC -1- PATHWAY: ISOPRENOID BIOSYNTHETIC PATHWAY WHOSE END PRODUCTS  
 CC INCLUDE DOLICHOLS, VITAMINS A, D, E, AND K, STEROID HORMONES,  
 CC CAROTENOIDS BILE ACIDS AND CHOLESTEROL.  
 CC -1- SUBCELLULAR LOCATION: Peroxisomal.  
 CC -1- SIMILARITY: BELONGS TO THE IPP ISOMERASE TYPE 1 FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AF003836; AAC53283.1;  
 DR InterPro; IPR002667; IPP\_isomerase.  
 DR InterPro; IPR000086; NUDIX\_hydrolase.  
 DR Pfam; PF00293; NUDIX; 1.  
 DR ProDom; PD004109; IPP\_isomerase; 1.  
 KW Isomerase; Isoprene biosynthesis; Cholesterol biosynthesis;  
 KW Sterol biosynthesis; peroxisome; Magnesium.

```
FT ACT_SITE 86 86 BY SIMILARITY.
FT ACT_SITE 148 148 BY SIMILARITY.
FT SITE 225 227 MICROBODY TARGETING SIGNAL.
SQ SEQUENCE 227 AA; 26317 MW; F500A6586385E803 CRC64;

Query Match 62.5%; Score 35; DB 1; Length 227;
Best Local Similarity 70.0%; Pred. No. 8.5;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMHY 10
DB 121 EEVDPMEMHY 130

RESULT 8
GAAB_METH STANDARD; PRT; 308 AA.
AC 026806;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE GMP synthase [glutamine-hydrolyzing] subunit B (EC 6.3.5.2) (GMP synthetase).
GN GUAAB OR MTH710.
OS Methanobacterium thermoautotrophicum.
OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
OC Methanobacteriaceae; Methanothermobacter.
OX NCBI_TaxID=187420;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Delta H;
RX MEDLINE=98037514; PubMed=9371463;
RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,
RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
RA Harrison D., Hoang L., Keagle P., Lumm W., Pothier B., Qiu D.,
RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
RA Jiwan N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,
RA McDougall S., Shimer G., Goyal A., Pietrowski S., Church G.M.,
RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.,
RA "Complete genome sequence of Methanobacterium thermoautotrophicum
RT Gelsht: functional analysis and comparative genomics.";
RL J. Bacteriol. 179:7135-7155(1997).
CC -1- CATALYTIC ACTIVITY: ATP + xanthosine 5'-phosphate + L-glutamine +
CC H(2)O = AMP + diphosphate + GMP + L-glutamate.
CC -1- PATHWAY: GMP biosynthesis.
CC -1- SUBUNIT: HETERODIMER COMPOSED OF A GLUTAMINE AMIDOTRANSFERASE
CC SUBUNIT (A) AND A GMP SYNTHASE SUBUNIT (B) (POTENTIAL).
CC -1- SIMILARITY: BELONGS TO THE GMP SYNTHASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AE000850; AAB85215.1; ALT_INIT.
CC HSPSP; P04079; IGPM.
CC InterPro: IPR001674; GMP_synt_C.
CC Pfam: PF00958; GMP_synt_C; 1.
CC TIGRFAMs: TIGR00884; guaA_Cterm; 1.
CC Ligase; GMP biosynthesis; Purine biosynthesis; ATP-binding;
KW Complete proteome. 184 GMP-BINDING (BY SIMILARITY).
FT DOMAIN 33 184
FT NP_BIND 29 35 ATP (BY SIMILARITY).
SQ SEQUENCE 308 AA; 34403 MW; F2DCF6ED202CAEC1 CRC64;

Query Match 62.5%; Score 35; DB 1; Length 308;
Best Local Similarity 63.6%; Pred. No. 12;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMHY 11
```

```
DB 216 EEVVEGLHES 226

RESULT 9
SLS1_YARLI STANDARD; PRT; 426 AA.
AC 099158;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE SLS1 protein precursor.
GN SLS1.
OS Yarrowia lipolytica (Candida lipolytica).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Dipodascaceae; Yarrowia.
OX NCBI_TaxID=4952;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 20460 / W29;
RX MEDLINE=96216076; PubMed=8662639;
RA Boisrame A., Beckerich J.-M., Gaillardin C.; is involved in the
RT protein translocation process in the yeast Yarrowia lipolytica.";
RL J. Biol. Chem. 271:11668-11675(1996).
CC -1- FUNCTION: INVOLVED IN THE PROTEIN TRANSLOCATION PROCESS. MAY
CC INTERACT DIRECTLY WITH TRANSLATING POLYPEPTIDES TO FACILITATE
CC THEIR TRANSFER AND/OR HELP THEIR FOLDING IN THE ER. IT IS NOT
CC REQUIRED FOR VIABILITY BUT IS ESSENTIAL FOR OPTIMAL GROWTH AT
CC ELEVATED TEMPERATURES.
CC -1- SUBCELLULAR LOCATION: Endoplasmic reticulum lumen.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: Z50154; CAA90516.1; -.
CC InterPro: IPR000886; ER_target.
CC PROSITE: PS00014; ER_TARGET; 1.
CC Endoplasmic reticulum; Signal.
FT SIGNAL 1 17 POTENTIAL.
FT CHAIN 18 426 SLS1 PROTEIN.
FT SITE 423 426 PREVENT SECRETION FROM ER (POTENTIAL).
SQ SEQUENCE 426 AA; 47201 MW; OACD7EF17540B8E2 CRC64;

Query Match 60.7%; Score 34; DB 1; Length 426;
Best Local Similarity 44.4%; Pred. No. 26;
Matches 4; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXGMH 9
DB 52 DQVIPAGLH 60

RESULT 10
PHSL_DESBA STANDARD; PRT; 513 AA.
AC PI3065;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Periplasmic [Nifese] hydrogenase large subunit (EC 1.18.99.1) (Nifese
OS hydrogenlyase large chain).
OS Desulfovibrio baculatus (Desulfomicrobium baculatus).
OX Bacteria; Proteobacteria; delta subdivision; Desulfomicrobium.
OX NCBI_TaxID=899;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88058744; PubMed=3316183;
```

Menon N.K., Peck H.D. Jr., le Gall J., Przybyla A.E.;  
 "Cloning and sequencing of the genes encoding the large and small  
 subunits of the periplasmic (Nifese) hydrogenase of *Desulfovibrio*  
*baculatus*."; J. Bacteriol. 169:5401-5407(1987).  
 [2]  
 REVISIONS.  
 Menon N.K., Peck H.D. Jr., le Gall J., Przybyla A.E.;  
 J. Bacteriol. 170:4429-4429(1988).  
 [3]  
 X-RAY CRYSTALLOGRAPHY (2.15 ANGSTROMS).  
 RX MEDLINE-99306038; PubMed-10378275;  
 RA Garcin E., Vernede X., Hachikian E.C., Volbeda A., Frey M.,  
 RA Fontecilla-Camps J.C.;  
 "The crystal structure of a reduced [NiFeSe] hydrogenase provides an  
 image of the activated catalytic center."; Structure 7:557-566(1999).  
 RT Structure 7:557-566(1999).  
 CC -1- CATALYTIC ACTIVITY: 2 reduced ferredoxin + 2 H(+) = 2 oxidized  
 CC ferredoxin + H(2).  
 CC -1- COFACTOR: ONE NICKEL ION, TWO 4FE-4S CLUSTERS AND ONE  
 CC SELENOCYSTEINE.  
 CC -1- SUBUNIT: HETERODIMER OF A LARGE AND A SMALL SUBUNIT.  
 CC -1- SUBCELLULAR LOCATION: Periplasmic.  
 CC -1- MISCELLANEOUS: PERHAPS THE LEADER OF THE SMALL SUBUNIT SERVES AS A  
 CC TRANSPORT VEHICLE FOR BOTH SUBUNITS.  
 CC -1- SIMILARITY: BELONGS TO THE [NIFE]/[NIFES] HYDROGENASE LARGE  
 CC SUBUNIT FAMILY.  
 -----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).  
 -----  
 EMBL; M18271; AAA23375.2; -;  
 DR PIR; A33101; HODVLB.  
 DR PDB; 1CC1; 01-JUN-99.  
 DR InterPro: IPR001501; Ni\_hdl.  
 DR Pfam; PF00374; Nifese\_Hases; 1.  
 DR PROSITE; PS00507; NI\_HGENASE\_L1; 1.  
 DR PROSITE; PS00508; NI\_HGENASE\_L2; 1.  
 KW Oxidoreductase; Periplasmic; Nickel; Selenium; Selenocysteine;  
 KW 3D-structure.  
 FT INIT\_MET 0 0  
 FT METAL 51 51 IRON 2.  
 FT METAL 70 70 NICKEL.  
 FT METAL 73 73 IRON 1 AND NICKEL.  
 FT METAL 444 444 IRON 2.  
 FT METAL 492 492 NICKEL.  
 FT METAL 495 495 IRON 1 AND NICKEL.  
 FT METAL 498 498 IRON 2.  
 FT SE\_CYS 492 492  
 SQ SEQUENCE 513 AA; 56683 MW; AC8285A6F80576FC CRC64;  
 Query Match 58.9%; Score 33; DB 1; Length 513;  
 Best Local Similarity 71.4%; Pred.No. 50;  
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 5 PYGMHYS 11  
 DB 297 PGGLHYS 303  
 -----  
 RESULT 11  
 MUTL\_BACSU  
 ID MUTL\_BACSU STANDARD; PRT; 627 AA.  
 AC P49850;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE DNA mismatch repair protein mutL.

GN  
 OS *Bacillus subtilis*.  
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
 OX NCBI\_TaxID=1423;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=168;  
 RX MEDLINE-96349107; PubMed-8760914;  
 RA Gineti F., Perego M., Albertini A.M., Galizzi A.;  
 "Bacillus subtilis mutL operon: identification, nucleotide  
 sequence and mutagenesis."; Microbiology 142:2021-2029(1996).  
 RT [2]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=168;  
 RX MEDLINE-98044033; PubMed-9384377;  
 RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,  
 Avevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,  
 Boriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,  
 Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,  
 Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,  
 Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,  
 Etian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,  
 Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,  
 Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,  
 Guisepi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,  
 Hilbert H., Holsappel S., Hosono S., Hulio M.F., Itaya M., Jones L.,  
 Joris B., Karamata D., Kasahara Y., Klaer-Blanchard M., Klein C.,  
 Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,  
 Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,  
 Lee S.M., Levine A., Liu H., Masuda S., Maue C., Medigue C.,  
 Medina N., Mellado R.P., Mizuno M., Moesti D., Nakai S., Noback M.,  
 Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,  
 Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,  
 Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,  
 Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadale I.,  
 Sato T., Scanlan E., Schleich S., Schroeter K., Scoffone F.,  
 Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,  
 Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,  
 Takeuchi M., Takamashi A., Tanaka T., Terpstra P., Tognoni A.,  
 Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,  
 Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,  
 Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,  
 Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.;  
 "The complete genome sequence of the Gram-positive bacterium *Bacillus*  
*subtilis*."; Nature 390:249-256(1997).  
 RT Nature 390:249-256(1997).  
 RL  
 CC -1- FUNCTION: THIS PROTEIN IS INVOLVED IN THE REPAIR OF MISMATCHES IN  
 CC DNA. IT IS REQUIRED FOR DAM-DEPENDENT METHYL-DIRECTED DNA MISMATCH  
 CC REPAIR. MAY ACT AS A "MOLECULAR MATCHMAKER", A PROTEIN THAT  
 CC PROMOTES THE FORMATION OF A STABLE COMPLEX BETWEEN TWO OR MORE  
 CC DNA-BINDING PROTEINS IN AN ATP-DEPENDENT MANNER WITHOUT ITSELF  
 CC BEING PART OF A FINAL EFFECTOR COMPLEX (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE DNA MISMATCH REPAIR MUTL/HEXB FAMILY.  
 -----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).  
 -----  
 EMBL; U27343; AAB19236.1; -;  
 DR EMBL; Z99112; CAB13578.1; -;  
 DR HSSP; P23367; 1BKN.  
 DR Subtilist; BG11402; mutL.  
 DR InterPro: IPR003594; ATPbind\_ATPase.  
 DR InterPro: IPR002099; DNA\_mis\_repair.  
 DR InterPro: IPR004359; HIS\_KIN\_sig.  
 DR Pfam; PF01119; DNA\_mis\_repair; 1.  
 DR Pfam; PF02518; HATPase\_c; 1.  
 DR TIGRFAMS; TIGR00585; mutL; 1.



DR PROSITE: PS00058; DNA\_MISMATCH\_REPAIR\_1; 1.  
KW DNA repair; Complete proteome.  
SQ SEQUENCE 627 AA; 70431 MW; 068A0509CC265343 CRC64;

Query Match 58.9%; Score 33; DB 1; Length 627;  
Best Local Similarity 54.5%; Pred. No. 61;  
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 EEVVPXGMHYS 11  
| : |||  
Db 488 ENIVPLTFHYS 498

RESULT 12  
KPBA\_CAEEL STANDARD; PRT; 1188 AA.  
AC P34335;  
DT 01-FEB-1994 (Rel. 28, Created)  
DT 15-JUN-2002 (Rel. 41, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Probable phosphorylase B kinase alpha regulatory chain (Phosphorylase  
DE kinase alpha subunit).  
GN C14B9.8.

OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
ON NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Bristol N2;  
RX MEDLINE=94150718; PubMed=7906398;  
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M., Coulson A.,  
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Fraser A.,  
RA Craxton M., Dear S., Du Z., Durbin R., Favell A., Fraser A.,  
RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,  
RA Johnston L., Jones M., Kershaw J., Kirsten J., Laisster N.,  
RA Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,  
RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Showkeen R.,  
RA Sims M., Smalton N., Smith A., Smith M., Sonhammer E., Staden R.,  
RA Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,  
RA Waterston R., Watson A., Weinstock L., Wilkinson-Sproat J.,  
RA Wohldman P.;  
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
RT elegans".  
RL Nature 368:32-38(1994).  
RN [2]  
RP REVISIONS.

RA Waterston R.;  
RA Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: PHOSPHORYLASE B KINASE CATALYZES THE PHOSPHORYLATION OF  
CC SERINE IN CERTAIN SUBSTRATES, INCLUDING TROPONIN I. THE ALPHA  
CC CHAIN MAY BIND CALMODULIN (BY SIMILARITY).  
CC -1- PATHWAY: Glycogen metabolism.  
CC -1- SIMILARITY: BELONGS TO THE PHOSPHORYLASE B KINASE REGULATORY  
CC CHAINS FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; L15188; AAA27954.2; -  
CC PIR; S44754; S44754.  
CC WormPep; C14B9.8; CE26870.  
CC Hypothetical protein; Glycogen metabolism; Calmodulin-binding.  
SQ SEQUENCE 1188 AA; 135558 MW; DE9BB875F3603863 CRC64;

Query Match 58.9%; Score 33; DB 1; Length 1188;  
Best Local Similarity 54.5%; Pred. No. 1.2e+02;  
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 EEVVPXGMHYS 11  
| : |||  
Db 950 EYAEEDGHIYS 960

RESULT 13  
VLTF\_BPT5 STANDARD; PRT; 1396 AA.  
ID P13390; O48502;  
AC 01-JAN-1990 (Rel. 13, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE L-shaped tail fiber protein (LTF protein).  
GN LTF.  
OS Bacteriophage T5.  
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae;  
OC T5-like viruses.  
OX NCBI\_TaxID=10726;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=95309401; PubMed=7789514;  
RA Kallman A.V., Kulshin V.E., Shlyapnikov M.G., Ksenzenko V.N.,  
RA Kryukov V.M.;  
RT "The nucleotide sequence of the bacteriophage T5 ltf gene".  
RL FEBS Lett. 366:46-48(1995).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Kallman A.V.;  
RA Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP PRELIMINARY PARTIAL SEQUENCE FROM N.A.  
RX MEDLINE=88289370; PubMed=3267228;  
RA Kallman A.V., Kryukov V.M., Bayev A.A.;  
RT "The nucleotide sequence of bacteriophage T5 DNA at the region  
RT between early and late genes".  
RL Nucleic Acids Res. 16:6230-6230(1988).  
CC -1- FUNCTION: NONESSENTIAL PROTEIN THAT MEDIATES BINDING TO THE  
CC POLYMANNOSE O ANTIGEN.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; X69460; CAA49220.1; -  
CC EMBL; AJ001191; CAA04591.1; -  
CC PIR; S01982; S01982.  
CC Late protein.  
CC CONFLICT 986 986 V -> A (IN REF. 2).  
SQ SEQUENCE 1396 AA; 147989 MW; 18CD2192F65FFFC1 CRC64;

Query Match 58.9%; Score 33; DB 1; Length 1396;  
Best Local Similarity 50.0%; Pred. No. 1.4e+02;  
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 EVVPXGMHYS 11  
: : | | | |  
Db 1360 KTIAPGDHYS 1369

RESULT 14  
TKNK\_BOVIN STANDARD; PRT; 126 AA.  
ID TKNK\_BOVIN  
AC P08858;  
DT 01-NOV-1988 (Rel. 09, Created)  
DT 01-NOV-1988 (Rel. 09, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Neurokinin B precursor (NKB) (Neuromedin K).  
GN TAC3 OR NKNB.

OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;  
 OC Bovidae; Bovinae; Bos.  
 OX NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=86313713; PubMed=3462746;  
 RA Kotani H., Hoshimaru M., Nawa H., Nakanishi S.;  
 RT "Structure and gene organization of bovine neuromedin K precursor.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 83:7074-7078(1986).  
 CC -1- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,  
 CC EVOKE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND  
 CC SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH  
 CC MUSCLES.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; M14351; AAA30723.1;  
 DR EMBL; M14347; AAA30723.1; JOINED.  
 DR EMBL; M14348; AAA30723.1; JOINED.  
 DR EMBL; M14349; AAA30723.1; JOINED.  
 DR EMBL; M14350; AAA30723.1; JOINED.  
 DR PIR; A25905; A25905.  
 DR InterPro; IPR003635; Neurokinin.  
 DR InterPro; IPR002040; Tachykinin.  
 DR ProDom; PD020370; Neurokinin; 1.  
 DR PROSITE; PS00267; TACHYKININ; 1.  
 DR Tachykinin; Neuropeptide; Cleavage on pair of basic residues;  
 KW Amidation; Signal.  
 FT SIGNAL 1 20 POTENTIAL.  
 FT PROPEP 21 83  
 FT PEPTIDE 86 95 NEUROKININ B.  
 FT PROPEP 99 126  
 FT MOD\_RES 95 95  
 FT MOD\_RES 95 95 AMIDATION (G-96 PROVIDE AMIDE GROUP).  
 SQ SEQUENCE 126 AA; 13871 MW; 446EF433498EC059 CRC64;  
 Query Match 57.1%; Score 32; DB 1; Length 126;  
 Best Local Similarity 66.7%; Pred. No. 19;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 EEVVPXGMH 9  
 Db 28 EQVPGGGH 36  
 ID RR2\_CHLVU STANDARD; PRT; 267 AA.  
 AC P56351;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE Chloroplast 30S ribosomal protein S2.  
 GN RPS2.  
 OS Chlorella vulgaris.  
 OG Chloroplast.  
 OC Eukaryota; Viridiplantae; Chlorophyta; Trebouxiophyceae; Chlorellales;  
 CC Chlorellaceae; Chlorella.  
 OX NCBI\_TaxID=3077;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=IAM C-27 / Tamiya;  
 RX MEDLINE=97303241; PubMed=9159184;  
 RA Wakasugi T., Nagai T., Kapoor M., Sugita M., Ito M., Ito S.,

RA Tsudzuki J., Nakashima K., Tsudzuki T., Suzuki Y., Hamada A., Ohta T.,  
 RA Inamura A., Yoshinaga K., Sugiura M.;  
 RT "Complete nucleotide sequence of the chloroplast genome from the  
 RT green alga Chlorella vulgaris: the existence of genes possibly  
 RT involved in chloroplast division.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 94:5967-5972(1997).  
 CC -1- SIMILARITY: BELONGS TO THE S2P FAMILY OF RIBOSOMAL PROTEINS.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; AB001684; BAA57862.1;  
 DR InterPro; IPR001865; Ribosomal\_S2.  
 DR Pfam; PF00318; Ribosomal\_S2; 1.  
 DR PRINTS; PR00395; RIBOSOMALS2.  
 DR TIGRfams; TIGR01011; rpsB\_bact; 1.  
 DR PROSITE; PS00962; RIBOSOMAL\_S2\_1; FALSE\_NEG.  
 DR PROSITE; PS00963; RIBOSOMAL\_S2\_2; 1.  
 DR Ribosomal protein; Chloroplast.  
 SQ SEQUENCE 267 AA; 30699 MW; 7903075340BD900F CRC64;  
 Query Match 57.1%; Score 32; DB 1; Length 267;  
 Best Local Similarity 40.0%; Pred. No. 40;  
 Matches 4; Conservative 4; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 EEVVPXGMHY 10  
 Db 8 EDMIQSGMHF 17  
 ID CGD2\_RAT STANDARD; PRT; 288 AA.  
 AC Q04827;  
 DT 01-OCT-1993 (Rel. 27, Created)  
 DT 01-OCT-1993 (Rel. 27, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Gl/S-specific cyclin D2 (Vin-1 proto-oncogene).  
 GN CCND2 OR VIN-1.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93275661; PubMed=8502486;  
 RA Hanna Z., Jankowski M., Tremblay P., Jiang X.M., Milatovich A.,  
 RA Francke U., Jolicoeur P.;  
 RT "The Vin-1 gene, identified by provirus insertional mutagenesis, is  
 RT the cyclin D2.";  
 RL Oncogene 8:1661-1666(1993).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95011623; PubMed=7926809;  
 RA Hosokawa Y., Onga T., Nakashima K.;  
 RT "Induction of D2 and D3 cyclin-encoding genes during promotion of the  
 RT Gl/S transition by prolactin in rat Nb2 cells.";  
 RL Gene 147:249-252(1994).  
 CC -1- FUNCTION: ESSENTIAL FOR THE CONTROL OF THE CELL CYCLE AT THE G1/S  
 CC (START) TRANSITION.  
 CC -1- SUBUNIT: INTERACTS WITH THE CDK4 AND CDK6 PROTEIN KINASES TO FORM  
 CC A SERINE/THREONINE KINASE HOLOENZYME COMPLEX. THE CYCLIN SUBUNIT  
 CC IMPARTS SUBSTRATE SPECIFICITY TO THE COMPLEX.  
 CC -1- SIMILARITY: BELONGS TO THE CYCLIN FAMILY. CYCLIN D SUBFAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

EMBL; L09752; AAA1010.1; -  
 EMBL; D16308; BAA03815.1; -  
 InterPro; IPR004366; Cyclin.  
 InterPro; IPR004367; Cyclin\_Cterm.  
 Pfam; PF00134; cyclin; 1.  
 Pfam; PF02984; cyclin\_C; 1.  
 SMART; SM00385; CYCLIN; 1.  
 PROSITE; PS00292; CYCLINS; 1.  
 Cyclin; Cell cycle; Cell division; Multigene family; Proto-oncogene.  
 CONFLICT 68 E -> G (IN REF. 2).  
 CONFLICT 104 C -> V (IN REF. 2).  
 CONFLICT 232 T -> A (IN REF. 2).  
 CONFLICT 288 AA; 32826 MW; 4B522BF4E9835FC1 CRC64;  
 SEQUENCE 288 AA; 32826 MW; 4B522BF4E9835FC1 CRC64;  
 Query Match 57.1%; Score 32; DB 1; Length 288;  
 Best Local Similarity 60.0%; Pred. No. 44;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMHY 10  
 III I I I  
 DB 73 EEVFPAMNY 82

-----

RESULT 17  
 CGD2\_HUMAN STANDARD; PRT; 289 AA.  
 ID CGD2\_HUMAN STANDARD; PRT; 289 AA.  
 AC P30279; Q13955; PubMed=1386336;  
 MEDLINE=92347851; PubMed=8455931;  
 RX Xiong Y., Menninger J., Beach D., Ward D.C.;  
 RA "Molecular cloning and chromosomal mapping of CCND genes encoding  
 RT human D-type cyclins."  
 RL Genomics 13:575-584(1992).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92347851; PubMed=8455931;  
 RA Palmero I., Holder A., Sinclair A.J., Dickson C., Peters G.;  
 RT "Cyclins D1 and D2 are differentially expressed in human B-lymphoid  
 RT cell lines."  
 RL Oncogene 8:1049-1054(1993).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Miyajima N.;  
 RL Submitted (MAR-1993) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-Bone marrow;  
 RA Strausberg R.;  
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
 RN [5]  
 RP SEQUENCE OF 1-240 FROM N.A.  
 RC TISSUE-Placenta;  
 RX MEDLINE=92347850; PubMed=1386335;  
 RA Inaba T., Matsushime H., Valentine M., Roussel M.F., Sherr C.J.,  
 RA Look A.T.;  
 RT "Genomic organization, chromosomal localization, and independent  
 RT expression of human cyclin D genes."  
 RL Genomics 13:565-574(1992).

-----

-1- FUNCTION: ESSENTIAL FOR THE CONTROL OF THE CELL CYCLE AT THE G1/S  
 (START) TRANSITION  
 -1- SUBUNIT: INTERACTS WITH THE CDK4 AND CDK6 PROTEIN KINASES TO FORM  
 A SERINE/THREONINE KINASE HOLOENZYME COMPLEX. THE CYCLIN SUBUNIT  
 IMPARTS SUBSTRATE SPECIFICITY TO THE COMPLEX.  
 -1- SIMILARITY: BELONGS TO THE CYCLIN FAMILY. CYCLIN D SUBFAMILY.  
 -----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

EMBL; M90813; AAA51926.1; -  
 EMBL; X68452; BAA04893.1; -  
 EMBL; D13639; BAA02802.1; -  
 EMBL; BC010958; AAI10958.1; -  
 EMBL; M80803; AAA51928.1; -  
 EMBL; M80800; AAA51928.1; JOINED.  
 EMBL; M80801; AAA51928.1; JOINED.  
 EMBL; M80802; AAA51928.1; JOINED.  
 PIR; A42822; A42822.  
 PIR; S26580; S26580.  
 Genew; HGNC:1583; CCND2.  
 MIM; 123833; -  
 InterPro; IPR004366; Cyclin.  
 InterPro; IPR004367; Cyclin\_Cterm.  
 Pfam; PF00134; cyclin; 1.  
 Pfam; PF02984; cyclin\_C; 1.  
 SMART; SM00385; CYCLIN; 1.  
 PROSITE; PS00292; CYCLINS; 1.  
 Cyclin; Cell cycle; Cell division; Multigene family.  
 CONFLICT 166 KL -> NV (IN REF. 5).  
 CONFLICT 224 T -> H (IN REF. 5).  
 SEQUENCE 289 AA; 33067 MW; E4E5FEF476D76D90 CRC64;

Query Match 57.1%; Score 32; DB 1; Length 289;  
 Best Local Similarity 60.0%; Pred. No. 44;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMHY 10  
 III I I I  
 DB 74 EEVFPAMNY 83

-----

RESULT 18  
 CGD2\_MOUSE STANDARD; PRT; 289 AA.  
 ID CGD2\_MOUSE STANDARD; PRT; 289 AA.  
 AC P30280;  
 DT 01-APR-1993 (Rel. 25, Created)  
 DT 01-APR-1993 (Rel. 25, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE G1/S-specific cyclin D2.  
 GN CCND2 OR CYL-2.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_Taxid=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92196134; PubMed=1372445;  
 RA Kiyokawa H., Busquets X., Powell C.T., Ngo L., Rifkind R.A.,  
 RA Marks P.A.;  
 RT "Cloning of a D-type cyclin from murine erythroleukemia cells."  
 RL Proc. Natl. Acad. Sci. U.S.A. 89:2444-2447(1992).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91235305; PubMed=1827757;  
 RA Matsushime H., Roussel M.F., Ashmun R.A., Sherr C.J.;  
 RT "Genomic organization, chromosomal localization, and independent  
 RT expression of human cyclin D genes."  
 RL Genomics 13:565-574(1992).

```

RL Cell 65:701-713(1991).
CC -!- FUNCTION: ESSENTIAL FOR THE CONTROL OF THE CELL CYCLE AT THE G1/S
CC (START) TRANSITION.
CC
CC -!- SUBUNIT: INTERACTS WITH THE CDK4 AND CDK6 PROTEIN KINASES TO FORM
CC A SERINE/THREONINE KINASE HOLOENZYME COMPLEX. THE CYCLIN SUBUNIT
CC IMPARTS SUBSTRATE SPECIFICITY TO THE COMPLEX.
CC
CC -!- SIMILARITY: BELONGS TO THE CYCLIN FAMILY. CYCLIN D SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M83749; AAA37519.1; -.
DR EMBL; M86182; AAA37503.1; -.
DR PIR; B40035; B40035.
DR PIR; A41984; A41984.
DR MGD; MGI:88314; Ccnd2.
DR InterPro; IPR004366; Cyclin.
DR InterPro; IPR004367; Cyclin_Cterm.
DR Pfam; PF00134; cyclin; 1.
DR Pfam; PF02984; cyclin_C; 1.
DR SMART; SM00385; CYCLIN; 1.
DR PROSITE; PS00292; CYCLINS; 1.
DR Cyclin; Cell cycle; Cell division.
DR KW Cyclin; Cell cycle; Cell division; Multigene family.
DR SQ SEQUENCE 289 AA; 32897 MW; 58F322771DD1DA3D CRC64;

Query Match 57.1%; Score 32; DB 1; Length 289;
Best Local Similarity 60.0%; Pred. No. 44;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMHY 10
Db ||| | | |
73 EEVFPPLAMNY 82

RESULT 19
CGDL_BRARE STANDARD; PRT; 291 AA.
AC Q90459;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE G1/S-specific cyclin D1.
GN CYCD1.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96138542; PubMed=8547308;
RA Yarden A., Salomon D., Geiger B.;
RT "Zebrafish cyclin D1 is differentially expressed during early
RL embryogenesis.";
RL Biochim. Biophys. Acta 1264:257-260(1995).
CC -!- FUNCTION: ESSENTIAL FOR THE CONTROL OF THE CELL CYCLE AT THE G1/S
CC (START) TRANSITION.
CC -!- SUBUNIT: INTERACTS WITH THE CDK4 AND CDK6 PROTEIN KINASES TO FORM
CC A SERINE/THREONINE KINASE HOLOENZYME COMPLEX. THE CYCLIN SUBUNIT
CC IMPARTS SUBSTRATE SPECIFICITY TO THE COMPLEX.
CC
CC -!- SIMILARITY: BELONGS TO THE CYCLIN FAMILY. CYCLIN D SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----

```

```

CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X87581; CAA60885.1; -.
DR ZFIN; ZDB-GENE-980526-176; Cycd1.
DR InterPro; IPR004366; Cyclin.
DR InterPro; IPR004367; Cyclin_Cterm.
DR Pfam; PF00134; cyclin; 1.
DR Pfam; PF02984; cyclin_C; 1.
DR SMART; SM00385; CYCLIN; 1.
DR PROSITE; PS00292; CYCLINS; 1.
DR Cyclin; Cell cycle; Cell division.
DR KW Cyclin; Cell cycle; Cell division.
DR SQ SEQUENCE 291 AA; 33067 MW; FA5274CB1B46D5EF CRC64;

Query Match 57.1%; Score 32; DB 1; Length 291;
Best Local Similarity 60.0%; Pred. No. 44;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMHY 10
Db ||| | | |
75 EEVFPPLAMNY 84

RESULT 20
CGDL_XENLA STANDARD; PRT; 291 AA.
ID CGDL_XENLA
AC P50755;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE G1/S-specific cyclin D1.
GN CCND1.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RA Cockrell M.J., Hunt T.;
RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: ESSENTIAL FOR THE CONTROL OF THE CELL CYCLE AT THE G1/S
CC (START) TRANSITION.
CC -!- SUBUNIT: INTERACTS WITH THE CDK4 AND CDK6 PROTEIN KINASES TO FORM
CC A SERINE/THREONINE KINASE HOLOENZYME COMPLEX. THE CYCLIN SUBUNIT
CC IMPARTS SUBSTRATE SPECIFICITY TO THE COMPLEX.
CC
CC -!- SIMILARITY: BELONGS TO THE CYCLIN FAMILY. CYCLIN D SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X89475; CAA61664.1; -.
DR InterPro; IPR004366; Cyclin.
DR InterPro; IPR004367; Cyclin_Cterm.
DR Pfam; PF00134; cyclin; 1.
DR Pfam; PF02984; cyclin_C; 1.
DR SMART; SM00385; CYCLIN; 1.
DR PROSITE; PS00292; CYCLINS; 1.
DR Cyclin; Cell cycle; Cell division.
DR KW Cyclin; Cell cycle; Cell division; Multigene family.
DR SQ SEQUENCE 291 AA; 32953 MW; AA747C5BD1679087 CRC64;

Query Match 57.1%; Score 32; DB 1; Length 291;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMHY 10
Db ||| | | |
73 EEVFPPLAMNY 82

```

```

RESULT 21
CGD2_CHICK STANDARD; PRT; 291 AA.
AC P49706;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE G1/S-specific cyclin D2.
GN CCND2.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96144302; PubMed=8566807;
RA Li H., Grenet J., Kidd V.J.;
RT "Structure and gene expression of avian cyclin D2.";
RL Gene 167:341-342(1995).
CC -!- FUNCTION: ESSENTIAL FOR THE CONTROL OF THE CELL CYCLE AT THE G1/S
CC (START) TRANSITION.
CC -!- SUBUNIT: INTERACTS WITH THE CDK4 AND CDK6 PROTEIN KINASES TO FORM
CC A SERINE/THREONINE KINASE HOLOENZYME COMPLEX. THE CYCLIN SUBUNIT
CC IMPARTS SUBSTRATE SPECIFICITY TO THE COMPLEX.
CC -!- SIMILARITY: BELONGS TO THE CYCLIN FAMILY. CYCLIN D SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X89476; CAA61665.1; -
DR EMBL; X83503; CAA58493.1; -
DR InterPro; IPR004366; Cyclin.
DR InterPro; IPR004367; Cyclin_Cterm.
DR Pfam; PF00134; cyclin; 1.
DR Pfam; PF02984; cyclin_C; 1.
DR SMART; SM00385; CYCLIN; 1.
DR PROSITE; PS00292; CYCLINS; 1.
KW Cyclin; Cell cycle; Cell division; Multigene family.
SQ SEQUENCE 291 AA; 32959 MW; 9A290F04F1531E89 CRC64;

Query Match 57.1%; Score 32; DB 1; Length 291;
Best Local Similarity 60.0%; Pred. No. 44;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMHY 10
Db ||||| |
74 EEVFPAMNY 83

RESULT 23
CGD1_CHICK STANDARD; PRT; 292 AA.
AC P55169;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE G1/S-specific cyclin D1.
GN CCND1.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
RA Li H., Lahti J.M., Kidd V.J.;
RT "FUNCTION: ESSENTIAL FOR THE CONTROL OF THE CELL CYCLE AT THE G1/S
CC (START) TRANSITION. INTERACTS WITH THE CDK4 AND CDK6 PROTEIN
CC KINASES (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE CYCLIN FAMILY. CYCLIN D SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

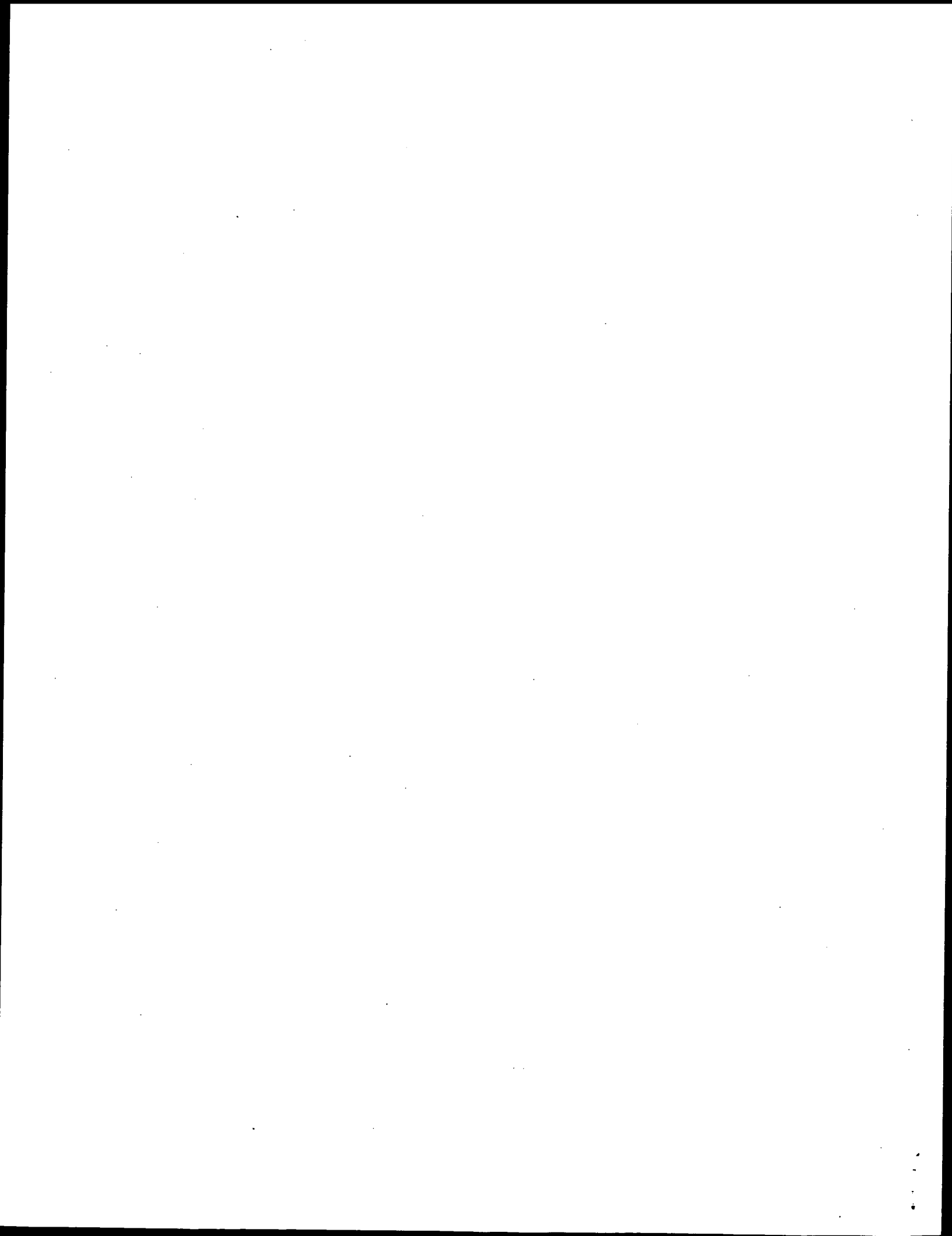
```

---

SEQUENCE FROM N.A.  
MEDLINE=94264323; PubMed=8204993;  
RX Rinkh R., Berger F., Bastard C., Klein B., French M., Archimbaud E.,  
RA Rouault J.-P., Santa Lucia B., Duret L., Vuillaume M.,  
RT "Rearrangement of CCND1 (BCL1/PRAD1) 3' untranslated region in  
mantle-cell lymphomas and t(11q13)-associated leukemias";  
RL Blood 83:3689-3696(1994).  
RN [6]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RA Strausberg R.;  
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
RN [7]  
RP INTERACTION WITH CDK4 AND CDK6.  
RX MEDLINE=94134440; PubMed=8302605;  
RA Bates S., Bonetta L., McAllan D., Parry D., Holder A., Dickson C.,  
RA Peters G.;  
RT "CDK6 (PLSTIRE) and CDK4 (PSK-J3) are a distinct subset of the  
cyclin-dependent kinases that associate with cyclin D1";  
RL Oncogene 9:71-79(1994).  
CC -!- FUNCTION: ESSENTIAL FOR THE CONTROL OF THE CELL CYCLE AT THE G1/S  
(START) TRANSITION.  
CC -!- SUBUNIT: INTERACTS WITH THE CDK4 AND CDK6 PROTEIN KINASES TO FORM  
A SERINE/THREONINE KINASE HOLOENZYME COMPLEX. THE CYCLIN SUBUNIT  
IMPARTS SUBSTRATE SPECIFICITY TO THE COMPLEX.  
CC -!- DISEASE: INVOLVED IN B-LYMPHOCTIC MALIGNANCY (PARTICULARLY  
MANTLE-CELL LYMPHOMA (MCL)) BY A CHROMOSOMAL TRANSLOCATION  
T(11;14)(Q13;Q32) THAT INVOLVES CCND1 AND IMMUNOGLOBULIN GENE  
REGIONS (BCL1 ONCOGENE). ACTIVATION OF CCND1 MAY BE ONCOGENIC BY  
DIRECTLY ALTERING PROGRESSION THROUGH THE CELL CYCLE.  
CC -!- DISEASE: INVOLVED IN A SUBSET OF PARATHYROID ADENOMAS BY A  
CHROMOSOMAL TRANSLOCATION T(11;11)(Q13;P15) THAT INVOLVES CCND1  
AND THE PARATHYROID HORMONE (PTH) ENHANCER (PRAD1 ONCOGENE).  
CC -!- SIMILARITY: BELONGS TO THE CYCLIN FAMILY. CYCLIN D SUBFAMILY.  
CC -!- DATABASE: NAME-Atlas Genet. Cytogenet. Oncol. Haematol.;  
WWW="http://www.infobiogen.fr/services/chroncancer/Genes/BCL1.html".  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; X59798; CAA42470.1; -;  
DR EMBL; M74092; -; NOT\_ANNOTATED\_CDS;  
DR EMBL; M64349; AAA52136.1; -;  
DR EMBL; M73554; AAA58392.1; -;  
DR EMBL; Z23022; CAA80558.1; -;  
DR EMBL; BC000076; AAH00076.1; -;  
DR EMBL; BC001501; AAH01501.1; -;  
DR EMBL; BC014078; AAH14078.1; -;  
DR PIR; B40268; B40268;  
DR PIR; S14794; S14794;  
DR PIR; A41523; A41523;  
DR PIR; A40034; A40034;  
DR PIR; S34295; S34295;  
DR Genew; HGNC:1582; CCND1.  
DR MIM; 151400; -;  
DR MIM; 168461; -;  
DR InterPro; IPR004366; Cyclin.  
DR InterPro; IPR004367; Cyclin\_Cterm.  
DR Pfam; PF00134; cyclin; 1.  
DR Pfam; PF02984; cyclin\_C; 1.  
DR SMART; SM00385; CYCLIN; 1.  
DR PROSITE; PS00292; CYCLINS; 1.  
KW Cyclin; Cell cycle; Cell division; Multigene family; Proto-oncogene;  
KW Chromosomal translocation.  
FT CONFLICT 130 130 N -> G (IN REF. 3).  
FT CONFLICT 168 169 MP -> IA (IN REF. 2).  
FT CONFLICT 188 188 L -> S (IN REF. 3).  
SQ SEQUENCE 295 AA; 33729 MW; 3CC00C9905F58D3A CRC64;

Query Match 57.1%; Score 32; DB 1; Length 295;  
Best Local Similarity 60.0%; Pred. No. 45;  
Matches 6; Conservative 1; Mismatches 3; Indels 0;  
QY 1 EEVVPXGMHY 10  
   | | | | | | | | | |  
DB 75 EEVFPPLANNY 84

Search completed: June 10, 2003, 13:40:21  
Job time : 5.5 secs





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2003, 13:30:25 ; Search time 25.7857 Seconds  
(without alignments)  
87.898 Million cell updates/sec

Title: US-09-909-164-11  
Perfect score: 56  
Sequence: 1 EFWVPXGMHYS 11

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 75 summaries

Database :

SPTREMBL\_21:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	38	67.9	1063	16 Q8RG86	Q8rg86 fusobacteri
2	37	66.1	308	16 Q9X2E2	Q9x2e2 thermotoga
3	37	66.1	322	17 Q9HLH8	Q9hlh8 thermoplas
4	37	66.1	1057	16 Q99UR5	Q99ur5 staphylococ
5	36	64.3	208	2 Q46486	Q46486 corynebacte
6	36	64.3	252	16 Q8YWP1	Q8ywp1 anabaena sp
7	36	64.3	819	10 Q9AVK4	Q9avk4 pisum sativ
8	35	62.5	139	2 Q57489	Q57489 bacteroides
9	35	62.5	156	16 Q9PC35	Q9pc35 xylella fas
10	35	62.5	233	10 Q40479	Q40479 nicotiana t
11	35	62.5	237	10 Q9LW50	Q9lw50 nicotiana s
12	35	62.5	317	9 Q38317	Q38317 lactobacill
13	35	62.5	425	5 Q9XVK4	Q9xvk4 caenorhabdi
14	35	62.5	510	10 Q9SA71	Q9sa71 arabidopsis
15	35	62.5	2042	17 Q8T207	Q8tz07 methanopyru
16	34	60.7	264	17 Q27902	Q27902 methanobact

## RESULT 1

Q8RG86  
ID Q8RG86  
AC Q8RG86;  
DT 01-JUN-2002 (TRENBLrel. 21, Created)  
DT 01-JUN-2002 (TRENBLrel. 21, Last sequence update)  
DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)  
DE Carbamoyl-phosphate synthase large chain (EC 6.3.5.5).  
GN FN0422.

## ALIGNMENTS

PRELIMINARY; PRT: 1063 AA.

17	34	60.7	279	16 Q9RXN9	Q9rxn9 deinococcus
18	34	60.7	285	16 Q98HU6	Q98hu6 rhizobium l
19	34	60.7	350	16 Q9RW92	Q9rw92 deinococcus
20	34	60.7	355	16 Q9XAM3	Q9xam3 streptomyce
21	34	60.7	360	17 Q27679	Q27679 methanobact
22	34	60.7	425	16 Q9KBA1	Q9kba1 bacillus ha
23	34	60.7	495	5 Q16912	Q16912 caenorhabdi
24	34	60.7	637	5 Q9Y0Y6	Q9y0y6 drosophila
25	34	60.7	678	12 Q9ELX6	Q9elx6 cercopithe
26	34	60.7	748	12 Q9YR01	Q9yr01 ranid herpe
27	34	60.7	1028	16 Q8XJ11	Q8xj11 bruceella me
28	34	60.7	2438	5 Q9VOL7	Q9vol7 drosophila
29	34	60.7	156	3 Q12479	Q12479 saccharomyc
30	33	58.9	216	16 Q9RDC1	Q9rdc1 streptomyce
31	33	58.9	257	8 Q99011	Q99011 prototheca
32	33	58.9	273	4 Q96N44	Q96n44 homo sapien
33	33	58.9	306	16 Q8XV07	Q8xv07 ralstonia s
34	33	58.9	367	16 Q916J1	Q916j1 pseudomonas
35	33	58.9	387	16 Q98FX1	Q98fx1 rhizobium l
36	33	58.9	426	3 Q9HG99	Q9hg99 mortierella
37	33	58.9	441	16 Q9KT98	Q9kt98 vibrio chol
38	33	58.9	466	16 Q9PKH7	Q9pkh7 chlamydia m
39	33	58.9	466	16 Q84218	Q84218 chlamydia t
40	33	58.9	534	17 Q29966	Q29966 archaeoglob
41	33	58.9	542	12 Q9WCW0	Q9wcw0 avian infec
42	33	58.9	545	10 Q9T025	Q9t025 arabidopsis
43	33	58.9	556	12 Q9QGT4	Q9qgt4 avian infec
44	33	58.9	558	12 Q9QGT3	Q9qgt3 avian infec
45	33	58.9	564	12 Q9QGT1	Q9qgt1 avian infec
46	33	58.9	565	12 Q9QGT5	Q9qgt5 avian infec
47	33	58.9	565	12 Q9QGT2	Q9qgt2 avian infec
48	33	58.9	567	5 Q9VEG2	Q9veg2 drosophila
49	33	58.9	572	4 Q8TEP5	Q8tep5 homo sapien
50	33	58.9	601	16 Q984F4	Q984f4 rhizobium l
51	33	58.9	716	11 P70521	P70521 rattus norv
52	33	58.9	796	10 Q8VY35	Q8vy35 zea mays (m
53	33	58.9	1442	17 Q96YH5	Q96yh5 sulfolobus
54	33	58.9	1471	4 Q8TEN9	Q8ten9 homo sapien
55	33	58.9	3472	1 Q74056	Q74056 cenarchaeum
56	32	57.1	103	11 Q9D0H9	Q9d0h9 mus musculu
57	32	57.1	153	5 Q76217	Q76217 anopheles g
58	32	57.1	153	13 P79919	P79919 xenopus lae
59	32	57.1	156	11 Q9D8L9	Q9dl9 mus musculu
60	32	57.1	190	13 Q57481	Q57481 sitostedio
61	32	57.1	191	11 Q99NB4	Q99nb4 rattus norv
62	32	57.1	197	16 Q9ZCF9	Q9zcf9 rickettsia
63	32	57.1	225	10 Q40129	Q40129 lycopersico
64	32	57.1	233	16 Q97166	Q97166 clostridium
65	32	57.1	235	5 Q9V5M0	Q9v5m0 drosophila
66	32	57.1	240	11 Q9DB09	Q9db09 mus musculu
67	32	57.1	270	16 Q9R9Q3	Q9r9q3 rhizobium m
68	32	57.1	291	13 Q8QFP4	Q8qfp4 brachydanio
69	32	57.1	322	17 Q978P6	Q978p6 thermoplas
70	32	57.1	329	5 Q8SV05	Q8svq5 encephalito
71	32	57.1	335	5 Q01456	Q01456 caenorhabdi
72	32	57.1	341	2 Q44548	Q44548 azotobacter
73	32	57.1	363	16 Q913T4	Q913t4 pseudomonas
74	32	57.1	370	2 Q9AG29	Q9ag29 pseudomonas
75	32	57.1	374	17 Q27916	Q27916 methanobact

```

OS Fusobacterium nucleatum (subsp. nucleatum).
OC Bacteria; Fusobacteria; Fusobacterium.
OX NCBI_TaxID=76856;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 25586;
RX MEDLINE=21886394; PubMed=11889109;
RA Kapatal V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A.,
RA Bhattacharyya A., Bartman A., Gardner W., Grechkin G., Zhu L.,
RA Vasileva O., Chu L., Kogan Y., Chaga O., Goltsman E., Bernal A.,
RA Larsen N., D'Souza M., Walunas T., Pusch G., Haselkorn R.,
RA Fonstein M., Kyripides N., Overbeek R.,
RT "Genome sequence and analysis of the oral bacterium Fusobacterium
nucleatum strain ATCC 25586."
EL J. Bacteriol. 184:2005-2018(2002).
DR EMBL; AE010534; AAL94625.1; -.
KW Ligase; Complete proteome.
SQ SEQUENCE 1063 AA; 118008 MW; 39700E10B7CCE411 CRC64;

Query Match 67.9%; Score 38; DB 16; Length 1063;
Best Local Similarity 60.0%; Pred. No. 48;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 EVVPXGMHY 11
Db ||| |::|
195 EIVPGLNYS 204

RESULT 2
Q9X2E2 PRELIMINARY; PRT; 308 AA.
AC Q9X2E2;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE F1SH protease activity modulator HFLK.
GN TM1822.
OS Thermotoga maritima.
OC Bacteria; Thermotogae; Thermotogales; Thermotogaceae; Thermotoga.
OX NCBI_TaxID=2336;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MSB8 / DSM 3109;
RX MEDLINE=99287316; PubMed=10360571;
RA Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
RA Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
RA McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
RA Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
RA Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
RT "Evidence for lateral gene transfer between Archaea and Bacteria from
genome sequence of Thermotoga maritima."
RL Nature 399:323-329(1999).
DR EMBL; AE001819; AAD36885.1; -.
DR TIGR; TM1822; -.
DR InterPro; IPR001107; Band_7.
DR InterPro; IPR003130; GED.
DR InterPro; IPR001972; StomatIn.
DR Pfam; PF01145; Band_7; 1.
DR PRINTS; PR00721; STOMATIN.
DR SMART; SM00302; GED; 1.
DR SMART; SM00244; PHB; 1.
KW Protease; Complete proteome.
SQ SEQUENCE 308 AA; 34778 MW; ADE03603E5101A9D CRC64;

Query Match 66.1%; Score 37; DB 16; Length 308;
Best Local Similarity 75.0%; Pred. No. 21;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMHY 10
Db ||| |::|
41 VVPSGHIY 48

us-09-909-164-11.rspt

RESULT 3
Q9HLH8 PRELIMINARY; PRT; 322 AA.
AC Q9HLH8;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE Glucose-fructose oxidoreductase related protein.
GN TR0250.
OS Thermoplasma acidophilum.
OC Archaea; Euryarchaeota; Thermoplasmata; Thermoplasmatales;
OC Thermoplasmataceae; Thermoplasma.
OX NCBI_TaxID=2303;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DSM 1728;
RX MEDLINE=20479972; PubMed=11029001;
RA Ruepp A., Graml W., Santos-Martinez M.-L., Koretke K.K., Volker C.,
RA Mewes H.-W., Frishman D., Stocker S., Lupas A.N., Baumeister W.;
RT "The genome sequence of the thermoacidophilic scavenger Thermoplasma
acidophilum."
RL Nature 407:508-513(2000).
DR EMBL; AL445063; CAC11395.1; -.
DR InterPro; IPR000683; GFO_IDH_MoCA.
DR Pfam; PF01408; GFO_IDH_MoCA; 1.
KW Complete proteome.
SQ SEQUENCE 322 AA; 36918 MW; B8C239E71009D167 CRC64;

Query Match 66.1%; Score 37; DB 17; Length 322;
Best Local Similarity 75.0%; Pred. No. 22;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMHY 10
Db ||| |::|
66 VVPDGLHY 73

RESULT 4
Q99UR5 PRELIMINARY; PRT; 1057 AA.
AC Q99UR5;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Carbamoyl-phosphate synthase large chain.
GN PYRAB OR SAV1203 OR SA1046.
OS Staphylococcus aureus (strain Mu50 / ATCC 700699), and
OS Staphylococcus aureus (Strain N315).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=158878; 158879;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=S. aureus (strain Mu50), and S. aureus (strain N315);
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
aureus."
RL Lancet 357:1225-1240(2001).
DR EMBL; AP003361; BAB57365.1; -.
DR EMBL; AP003132; BAB42298.1; -.
DR HSSP; P00968; ICS0.
DR InterPro; IPR000901; CPSase.
DR InterPro; IPR004362; MGS_like.
DR InterPro; IPR000169; SHprot_acsfe.
DR Pfam; PF00289; CPSase_L_chain; 2.

```

DR Pfam; PF02786; CPSase\_L\_D2; 2.  
 DR Pfam; PF02787; CPSase\_L\_D3; 1.  
 DR Pfam; PF02142; MGS; 1.  
 DR PRINTS; PR00098; CPSASE.  
 DR PROSITE; PS00866; CPSASE\_1; 2.  
 DR PROSITE; PS00867; CPSASE\_2; UNKNOWN\_2.  
 DR PROSITE; PS00639; THIOL\_PROTEASE\_HIS; UNKNOWN\_1.  
 KW Complete proteome.  
 SQ SEQUENCE 1057 AA; 117171 MW; E3E179EF0591F0F8 CRC64;

Query Match 66.18; Score 37; DB 16; Length 1057;  
 Best Local Similarity 60.08; Pred. No. 77;  
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 EVVPXGMHYS 11  
 : : : : :  
 Db 190 EIVSNGLHYS 199  
 : : : : :  
 : : : : :

RESULT 5  
 Q46486 PRELIMINARY; PRT; 208 AA.  
 ID Q46486  
 AC Q46486  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)  
 DE Hypothetical 23.0 kDa protein (GcrA).  
 GN GCR.  
 OS Corynebacterium xerosis, and  
 OS Corynebacterium striatum.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Corynebacterineae; Corynebacteriaceae;  
 OC Corynebacterium.  
 OX NCBI\_TaxID=1725, 43770;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=C.xerosis; STRAIN=M82B;  
 RX MEDLINE=96117603; PubMed=8559800;  
 RA Tauch A., Kassing F., Kalinowski J., Puhler A.;  
 RT "The Corynebacterium xerosis composite transposon Tn5432 consists of  
 RT two identical insertion sequences, designated IS1249, flanking the  
 RT erythromycin resistance gene ermCX.";  
 RL Plasmid 34:119-131(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=C.striatum; STRAIN=M82B;  
 RX MEDLINE=20194806; PubMed=10732668;  
 RA Tauch A., Krief S., Kalinowski J., Puhler A.;  
 RT "The 51,409-bp R-plasmid pTP10 from the multidrug-resistant clinical  
 RT isolate Corynebacterium striatum M82B is composed of DNA segments  
 RT initially identified in soil bacteria and in plant, animal, and human  
 RT pathogens.";  
 RL Mol. Gen. Genet. 263:1-11(2000).  
 DR EMBL; U21300; AAC95478.1; -;  
 DR EMBL; AF024666; AAG03390.1; -;  
 KW Hypothetical protein; Plasmid.  
 SQ SEQUENCE 208 AA; 23012 MW; F1504BE1ECDEB5A6 CRC64;

Query Match 64.38; Score 36; DB 2; Length 208;  
 Best Local Similarity 50.08; Pred. No. 22;  
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 EVVPXGMHYS 11  
 : : : : :  
 Db 130 DVIPEGHYA 139  
 : : : : :  
 : : : : :

RESULT 6  
 Q8YWP1 PRELIMINARY; PRT; 252 AA.  
 ID Q8YWP1  
 AC Q8YWP1  
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)

DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 DE Hypothetical protein ALR1563.  
 GN ALR1563.  
 OS Anabaena sp. (strain PCC 7120).  
 OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.  
 OX NCBI\_TaxID=103690;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21595285; PubMed=11759840;  
 RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,  
 RA Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,  
 RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,  
 RA Nakazaki N., Shimpo S., Sugimoto M., Takazawa M., Yamada M.,  
 RA Yasuda M., Tabata S.;  
 RT "Complete genomic sequence of the filamentous nitrogen-fixing  
 RT cyanobacterium Anabaena sp. strain PCC 7120.";  
 RL DNA Res. 8:205-213(2001).  
 DR EMBL; AF003586; BAB77929.1; -;  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 252 AA; 28831 MW; 925572DA5D1CA519 CRC64;

Query Match 64.38; Score 36; DB 16; Length 252;  
 Best Local Similarity 50.08; Pred. No. 27;  
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMHY 10  
 : : : : :  
 Db 235 EMIVPAGLHF 244  
 : : : : :  
 : : : : :

RESULT 7  
 Q9AVK4 PRELIMINARY; PRT; 819 AA.  
 ID Q9AVK4  
 AC Q9AVK4  
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE SCARECROW.  
 GN PSSCR.  
 OS Pisum sativum (Garden pea).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Viciae; Pisum.  
 OX NCBI\_TaxID=3888;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. ALASKA;  
 RX MEDLINE=21231727; PubMed=11333309;  
 RA Sassa N., Matsushita Y., Nakamura T., Nyunoya H.;  
 RT "The Molecular Characterization and in situ Expression Pattern of Pea  
 RT SCARECROW Gene.";  
 RL Plant Cell Physiol. 42:385-394(2001).  
 DR EMBL; AB048713; BAB39155.1; -;  
 DR InterPro; IPR001444; Flag\_bb\_rod.  
 DR InterPro; IPR005202; GRAS.  
 DR Pfam; PF03514; GRAS; 1.  
 DR PROSITE; PS00588; FLAGELLA\_BB\_ROD; UNKNOWN\_1.  
 SQ SEQUENCE 819 AA; 90372 MW; 41B67BD6DC72ADFA CRC64;

Query Match 64.38; Score 36; DB 10; Length 819;  
 Best Local Similarity 45.58; Pred. No. 94;  
 Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMHYS 11  
 : : : : :  
 Db 343 DDVVPSTLHFS 353  
 : : : : :  
 : : : : :

RESULT 8  
 Q57489 PRELIMINARY; PRT; 139 AA.  
 ID Q57489  
 AC Q57489  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)

DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)  
 DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)  
 DE DNA ligase (Fragment).  
 OS Bacteroides nodosus (Dichelobacter nodosus).  
 OC Bacteria; Proteobacteria; gamma subdivision; Cardiobacteriaceae;  
 OC Dichelobacter.  
 OX NCBI\_TaxID=870;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96020672; PubMed=7476204;  
 RA Rood J.I.;  
 RA Rood J.I.;  
 RA "A multiple site-specific DNA-inversion model for the control of Omp1  
 RT phase and antigenic variation in Dichelobacter nodosus.";  
 RL Mol. Microbiol. 17:183-196(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96257263; PubMed=8654969;  
 RA Billington S.J., Sinistaj M., Cheetham B.F., Ayres A., Moses E.K.,  
 RA Katz M.E., Rood J.I.;  
 RT Identification of a native Dichelobacter nodosus plasmid and  
 RT implications for the evolution of the vap regions.";  
 RL Gene 172:111-116(1996).  
 DR EMBL; U02462; AAB12366.1; -  
 DR InterPro: IPR001357; BRCT.  
 DR Pfam: PF00533; BRCT; 1.  
 DR SMART; SM00292; BRCT; 1.  
 DR PROSITE; PS00172; BRCT; 1.  
 KW Ligase.  
 FT NON\_TER  
 SQ SEQUENCE 139 AA; 15052 MW; E0E110AA4B7D4708 CRC64;  
 Query Match 62.5%; Score 35; DB 2; Length 139;  
 Best Local Similarity 55.6%; Pred. No. 24;  
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
 QY 3 VVPXGMHYS 11  
 DB 21 IVPAGVHWS 29  
 RESULT 9  
 Q9PC35 PRELIMINARY; PRT; 156 AA.  
 AC Q9PC35  
 DT 01-OCT-2000 (TReMBLrel. 15, Created)  
 DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)  
 DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)  
 DE Hypothetical protein Xf1950.  
 GN Xf1950  
 OS Xylella fastidiosa.  
 OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;  
 OC Xylella.  
 OX NCBI\_TaxID=2371;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=9A5C;  
 RX MEDLINE=20365717; PubMed=10910347;  
 RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,  
 RA Alvarenga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,  
 RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,  
 RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carrier H.,  
 RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,  
 RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,  
 RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,  
 RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,  
 RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,  
 RA Ho P.L., Hoheisel J.D., Junqueira M.L., Kemper E.B., Kitajima J.P.,  
 RA Krieger J.E., Kuramae E.V., Laigret F., Lambais M.R., Leite J.C.C.,  
 RA Lemos E.G.M., Lemos M.F.F., Lopes S.A., Lopes C.R., Machado J.A.,  
 RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,  
 RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,  
 RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,

RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,  
 RA Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,  
 RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,  
 RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,  
 RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,  
 RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,  
 RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,  
 RA da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,  
 RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsubako M.H.,  
 RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,  
 RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;  
 RT "The genome sequence of the plant pathogen Xylella fastidiosa.";  
 RL Nature 406:151-159(2000).  
 DR EMBL; AE004014; AAF64752.1; -  
 DR InterPro: IPR002545; Chew.  
 DR SMART; SM00260; Chew; 1.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 156 AA; 17144 MW; D8358619C6671A5D CRC64;  
 Query Match 62.5%; Score 35; DB 16; Length 156;  
 Best Local Similarity 55.6%; Pred. No. 27;  
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 EEVVPXGMH 9  
 DB 119 EEILPOGVH 127  
 RESULT 10  
 Q40479 PRELIMINARY; PRT; 233 AA.  
 AC Q40479  
 DT 01-NOV-1996 (TReMBLrel. 01, Created)  
 DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)  
 DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)  
 DE EXEBP-2.  
 OS Nicotiana tabacum (Common tobacco).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.  
 OX NCBI\_TaxID=4097;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BY4; TISSUE=LEAF;  
 RX MEDLINE=95276459; PubMed=7756828;  
 RA Ohme-Takagi M., Shinshi H.;  
 RT "Ethylene-inducible DNA binding proteins that interact with an  
 RT ethylene responsive element.";  
 RL Plant Cell 7:173-182(1995).  
 DR EMBL; D38126; BAA07324.1; -  
 DR HSSP; O80337; 2GCC.  
 DR TRANSFAC; T02654; -  
 DR InterPro: IPR001471; TF\_ERF.  
 DR Pfam; PF00847; AP2-domain; 1.  
 DR PRINTS; PR00367; THRSPLEMT.  
 DR PRODOM; PD001423; TF\_AP2; 1.  
 DR SMART; SM00380; AP2; 1.  
 SQ SEQUENCE 233 AA; 25563 MW; 6CD16783582C0CB5 CRC64;  
 Query Match 62.5%; Score 35; DB 10; Length 233;  
 Best Local Similarity 60.0%; Pred. No. 41;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 EEVVPXGMH 10  
 DB 90 QAVVPKGRHY 99  
 RESULT 11  
 Q9LW50 PRELIMINARY; PRT; 237 AA.  
 ID Q9LW50  
 AC Q9LW50  
 DT 01-OCT-2000 (TReMBLrel. 15, Created)

DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE Ethylene-responsive element binding factor.  
 GN NSERF2.  
 OS Nicotiana sylvestris (Wood tobacco).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.  
 OX NCBI\_TaxID=4096;  
 [1]  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE=20399450; PubMed=1045353;  
 RA Kitajima S., Koyama T., Ohme-Takagi M., Shinshi H., Sato F.;  
 RT "Characterization of gene expression of NSERFs, transcription factors  
 of basic PR genes from Nicotiana sylvestris."  
 RL Plant Cell Physiol. 41:817-824(2000).  
 DR EMBL; AB016264; BAA97122.1; -.  
 DR HSP; O80337; 2GCC.  
 DR InterPro: IPR001471; TF\_ERF.  
 DR Pfam; PF00847; AP2-domain; 1.  
 DR PRINTS; PR00367; ETHRSPELEMT.  
 DR ProDom; PD001423; TF\_AP2; 1.  
 DR SMART; SM00380; AP2; 1.  
 SQ SEQUENCE 237 AA; 26243 MW; 01BC3EEB51E46298 CRC64;  
 Query Match 62.5%; Score 35; DB 10; Length 237;  
 Best Local Similarity 60.0%; Pred. No. 41;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 EEVVPXGMHY 10  
 : : : : :  
 Db 94 QAVVPKGRHY 103  
 : : : : :  
 RESULT 12  
 Q38317 PRELIMINARY; PRT; 317 AA.  
 AC Q38317;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE Lysin.  
 GN LYS.  
 OS Lactobacillus bacteriophage phi adh.  
 OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae.  
 OX NCBI\_TaxID=12417;  
 [1]  
 RN SEQUENCE FROM N.A.  
 RA Altermann E.;  
 RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.  
 [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99384014; PubMed=10452953;  
 RA Altermann E., Klein J., Henrich B.;  
 RT "Primary structure and features of the genome of the Lactobacillus  
 gasserii temperate bacteriophage phi-adh."  
 RL Gene 236:333-346(1999).  
 [3]  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE=95138034; PubMed=7836307;  
 RA Henrich B., Binshofer B., Blaesi U.;  
 RT "Primary structure and functional analysis of the lysis genes of  
 Lactobacillus gasserii bacteriophage phi-adh."  
 RL J. Bacteriol. 177:723-732(1995).  
 [4]  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE=93231538; PubMed=8472961;  
 RA Fremaux C., De Antoni G., Raya R., Klaenhammer T.;  
 RT "Genetic organization and sequence of the region encoding integrative  
 functions from Lactobacillus gasserii temperate bacteriophage phi-  
 adh."  
 RL Gene 126:61-66(1993).  
 [5]

RP SEQUENCE FROM N.A.  
 RA Engel G., Altermann E., Klein J., Henrich B.;  
 RT "Structure of a genome region of the Lactobacillus gasserii temperate  
 phage phi adh covering a repressor gene and cognate promoters."  
 RL Gene 210:67-70(1998).  
 DR EMBL; AJ131519; CAB52540.1; -.  
 DR InterPro: IPR002053; GH\_25.  
 DR InterPro: IPR003646; SH3\_bac.  
 DR Pfam; PF01183; Glyco\_hydro\_25; 1.  
 DR ProDom; PD004620; GH\_25; 1.  
 DR SMART; SM00287; SH3B; 1.  
 SQ SEQUENCE 317 AA; 34703 MW; 9FF2715EEA3561C7 CRC64;  
 Query Match 62.5%; Score 35; DB 9; Length 317;  
 Best Local Similarity 66.7%; Pred. No. 56;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 QY 3 VVPXGMHYS 11  
 : : : : :  
 Db 60 VVPMGYHYA 68  
 : : : : :  
 RESULT 13  
 Q9XVK4 PRELIMINARY; PRT; 425 AA.  
 AC Q9XVK4;  
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
 DT 01-WAR-2002 (TrEMBLrel. 20, Last annotation update)  
 DE R10D12.10 protein.  
 GN R10D12.10.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 [1]  
 RN SEQUENCE FROM N.A.  
 RA Percy C.M.;  
 RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.  
 [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99069613; PubMed=9851916;  
 RA none;  
 RT "Genome sequence of the nematode C.elegans: A platform for  
 investigating biology."  
 RL Science 282:2012-2018(1998).  
 DR EMBL; Z81109; CAB03241.1; -.  
 DR InterPro: IPR000719; Euk\_pkinase.  
 DR Pfam; PF00069; pkinase; 1.  
 DR ProDom; PD000001; Euk\_pkinase; 1.  
 DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
 KW ATP-binding; Transferase.  
 SQ SEQUENCE 425 AA; 49410 MW; 5D96E29B08C8E9D6 CRC64;  
 Query Match 62.5%; Score 35; DB 5; Length 425;  
 Best Local Similarity 50.0%; Pred. No. 76;  
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 EEVVPXGMHY 10  
 : : : : :  
 Db 335 EQVPGGLQY 344  
 : : : : :  
 RESULT 14  
 Q9SA71 PRELIMINARY; PRT; 510 AA.  
 AC Q9SA71;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE TS18.1 protein.  
 GN TS18.1.  
 OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. COLUMBIA;  
 RA Vysotskaia V.S.; Schwartz J.R.; Yu G.; Toriumi M.; Lenz C.; Liu S.;  
 RA Li J.; Kremenetskaia I.; Luros J.; Ngan I.; Gonzalez A.; Altafi H.;  
 RA Araujo R.; Chao Q.; Conn L.; Conway A.B.; Dunn P.; Hansen N.;  
 RA Hulzar L.; Kim C.; Palm C.; Rowley D.; Shinn P.; Walker M.;  
 RA Davis R.W.; Ecker J.R.; Federspiel N.A.; Theologis A.;  
 RT "Arabidopsis thaliana chromosome 1 BAC T518 sequence."  
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).  
 DR EMBL; AC007060; AAD25743.1; -;  
 KW Transmembrane.  
 SQ SEQUENCE 510 AA; 55089 MW; 0DC63CC41F4712F8 CRC64;

Query Match 62.5%; Score 35; DB 10; Length 510;  
 Best Local Similarity 60.0%; Pred. No. 92;  
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 EVVVPXGMHY 10  
 ||| | | | |  
 DB 12 EEVKKPGIHF 21

RESULT 15  
 Q8TZ07 PRELIMINARY; PRT; 2042 AA.  
 AC Q8TZ07;  
 DT 01-JUN-2002 (TRENBLrel. 21, Created)  
 DT 01-JUN-2002 (TRENBLrel. 21, Last sequence update)  
 DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)  
 DE Predicted protein of the CobN/Mg-chelatase family.  
 GN MK0134.  
 OS Methanopyrus kandleri.  
 OC Archaea; Euryarchaeota; Methanopyri; Methanopyrales; Methanopyraceae;  
 OC Methanopyrus.  
 OX NCBI\_TaxID=2320;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=AV19 / DSM 6324 / JCM 9639;  
 RX MEDLINE=21927647; PubMed=11930014;  
 RA Slesarev A.I.; Mezhevaya K.V.; Makarova K.S.; Polushin N.N.;  
 RA Shcherbinina O.V.; Shakhova V.V.; Belova G.I.; Aravind L.;  
 RA Natale D.A.; Rogozin I.B.; Tatusov R.L.; Wolf Y.I.; Stetter K.O.;  
 RA Malykh A.G.; Koonin E.V.; Kozlyavkin S.A.;  
 RT "The complete genome of hyperthermophile Methanopyrus kandleri AV19  
 RL and monophyly of archaeal methanogens."  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:4644-4649(2002).  
 DR EMBL; AE010313; AA001351.1; -;  
 KW Complete proteome.  
 SQ SEQUENCE 2042 AA; 227716 MW; 83B6E092A62C112A CRC64;

Query Match 62.5%; Score 35; DB 17; Length 2042;  
 Best Local Similarity 75.08%; Pred. No. 4e+02;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 EVVVPXGMH 9  
 ||| | | | |  
 DB 924 EVVPIGLH 931

RESULT 16  
 O27902 PRELIMINARY; PRT; 264 AA.  
 AC O27902;  
 DT 01-JAN-1998 (TRENBLrel. 05, Created)  
 DT 01-JAN-1998 (TRENBLrel. 05, Last sequence update)  
 DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)  
 DE Diphthine synthase.

GN MTH1874.  
 OS Methanobacterium thermoautotrophicum.  
 OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;  
 OC Methanobacteriaceae; Methanothermobacter.  
 OX NCBI\_TaxID=187420;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=DELTA H;  
 RX MEDLINE=98037514; PubMed=9371463;  
 RA Smith D.R.; Doucette-Stamm L.A.; DeLoughery C.; Lee H.-M.; Dubois J.;  
 RA Aldredge T.; Bashirzadeh R.; Blakely D.; Cook R.; Gilbert K.;  
 RA Harrison D.; Hoang L.; Keagle P.; Lumm W.; Pothier B.; Qiu D.;  
 RA Spadafora R.; Vicare R.; Wang Y.; Wierzbowski J.; Gibson R.;  
 RA Jivani N.; Caruso A.; Bush D.; Safer H.; Patwell D.; Prabhakar S.;  
 RA McDougall S.; Shimer G.; Goyal A.; Pietrowski S.; Church G.M.;  
 RA Daniels C.J.; Mao J.-I.; Rice P.; Noelling J.; Reeve J.N.;  
 RT "Complete genome sequence of Methanobacterium thermoautotrophicum  
 RL J. Bacteriol. 179:7135-7155(1997).  
 DR EMBL; AE000940; AAB86340.1; -;  
 DR InterPro; IPR000878; CorPor\_Mettransf.  
 DR InterPro; IPR004551; Dphthm\_synthase.  
 DR Pfam; PF00590; TP\_methylase; 1.  
 DR TIGRFAMs; TIGR00522; dph5; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 264 AA; 28858 MW; 366BAE4E4D992C21 CRC64;

Query Match 60.7%; Score 34; DB 17; Length 264;  
 Best Local Similarity 62.5%; Pred. No. 74;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMHY 10  
 ||| | | | |  
 DB 235 VVPAGLHF 242

RESULT 17  
 Q9RXN9 PRELIMINARY; PRT; 279 AA.  
 AC Q9RXN9;  
 DT 01-MAY-2000 (TRENBLrel. 13, Created)  
 DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)  
 DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)  
 DE Hypothetical protein DR0271.  
 GN DR0271.  
 OS Deinococcus radiodurans  
 OC Bacteria; Thermus/Deinococcus group; Deinococci; Deinococcales;  
 OC Deinococcaceae; Deinococcus.  
 OX NCBI\_TaxID=1299;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=R1;  
 RX MEDLINE=20036896; PubMed=10567266;  
 RA White O.; Eisen J.A.; Heidelberg J.F.; Hickey E.K.; Peterson J.D.;  
 RA Dodson R.J.; Haft D.H.; Gwinn M.L.; Nelson W.C.; Richardson D.L.;  
 RA Moffat K.S.; Qin H.; Jiang L.; Pamphile W.; Crosby M.; Shen M.;  
 RA Vamathevan J.J.; Lam P.; McDonald L.; Utterback T.; Zaleski C.;  
 RA Makarova K.S.; Aravind L.; Daly M.J.; Minton K.W.; Fleischmann R.D.;  
 RA Ketchum K.A.; Nelson K.E.; Salzberg S.; Smith H.O.; Venter J.C.;  
 RA Fraser C.M.;  
 RT "Genome sequence of the radioresistant bacterium Deinococcus  
 RL radiodurans R1."  
 RL Science 286:1571-1577(1999).  
 DR EMBL; AE001889; AAF09867.1; -;  
 DR TIGR; DR0271; -;  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 279 AA; 31140 MW; DCEA100E0AE8831 CRC64;

Query Match 60.7%; Score 34; DB 16; Length 279;  
 Best Local Similarity 75.08%; Pred. No. 79;  
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 VPXGMHYS 11

```

RL Science 286:1571-1577(1999).
DR EMBL; AE001932; AAF10353.1; -.
DR TIGR; P07547; IDOS.
DR InterPro; IPR002658; DHQ_synthase.
DR Pfam; PF01761; DHQ_synthase; 1.
KW Complete proteome.
SQ SEQUENCE 350 AA; 37121 MW; 37601D08B2FB6116 CRC64;

Query Match 60.7%; Score 34; DB 16; Length 350;
Best Local Similarity 60.0%; Pred. No. 1e+02; 3; Indels 0; Gaps 0;
Matches 6; Conservative 1; Mismatches 3;

QY 2 EVVPGMHYS 11
DB 252 EAVAYGMHYA 261

RESULT 20
Q9XAM3 PRELIMINARY; PRT; 355 AA.
AC Q9XAM3
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Putative DNA ligase.
GN SC06707 OR SC4C6.17C.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Seeger S., Harris D.;
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA James K.D., Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RX MEDLINE=97000351; PubMed=8843436;
RA Redenbach M., Kieser H.M., Denapate D., Eichner A., Cullum J.,
RA Kinashi H., Hopwood D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL Mol. Microbiol. 21:77-96(1996).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neil S.,
RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2).";
RL Nature 417:141-147(2002).
DR EMBL; AL079355; CAB45581.1; -.
DR InterPro; IPR000977; DNA_ligase.
DR Pfam; PF01068; DNA_ligase; 1.
DR PROSITE; PS00697; DNA_LIGASE_A1; UNKNOWN_1.
KW Ligase.
SQ SEQUENCE 355 AA; 39873 MW; 7A995A55116077B1 CRC64;

Query Match 60.7%; Score 34; DB 16; Length 355;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 VVPGMHYS 11
DB 194 VVPGMHYS 202

RESULT 19
Q9RW92 PRELIMINARY; PRT; 350 AA.
AC Q9RW92
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE 3-dehydroquinase synthase.
GN DR0777
OS Deinococcus radiodurans.
OC Bacteria; Thermus/Deinococcus group; Deinococci; Deinococcales;
OC Deinococcaceae; Deinococcus.
OX NCBI_TaxID=1299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RL;
RX MEDLINE=20036896; PubMed=10567266;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zalewski C.,
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RT "Genome sequence of the radioresistant bacterium Deinococcus
RT radiodurans RL.";

```

Best Local Similarity 71.4%; Pred. No. 1e+02;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 VPXGMHY 10  
Db 20 IPPGMHY 26

## RESULT 21

O27679  
ID O27679 PRELIMINARY; PRT; 360 AA.  
AC O27679  
DT 01-JAN-1998 (TReMBLrel. 05, Created)  
DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)  
DE 01-JUN-2002 (TReMBLrel. 21, Last annotation update)  
DE Cell division protein.  
GN MTH1642.  
OS Methanobacterium thermoautotrophicum.  
OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;  
OC Methanobacteriaceae; Methanothermobacter.  
OX NCBI\_TaxID=187420;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=DELTA H;  
RX MEDLINE=98037514; PubMed=9371463;  
RA Smith D.R., Doucette-Stamm L.A., DeLoughery C., Lee H.-M., Dubois J.,  
RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,  
RA Harrison D., Hoang L., Keagle P., Lumm W., Pothier B., Qiu D.,  
RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,  
RA Jiwan N., Caruso A., Bush D., Goyal A., Pietrowski S., Prabhakar S.,  
RA McDougall S., Shimer G., Goyal A., Pietrowski S., Church G.M.,  
RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;  
RT "Complete genome sequence of Methanobacterium thermoautotrophicum  
delta: functional analysis and comparative genomics."  
RL J. Bacteriol. 179:7133-7155(1997).  
DR EMBL; AE000923; AAB6115.1;  
DR InterPro; IPR005140; eRF1.1.  
DR InterPro; IPR005141; eRF1.2.  
DR InterPro; IPR005142; eRF1.3.  
DR InterPro; IPR004405; eRF1.  
DR Pfam; PF03463; eRF1\_1; 1.  
DR Pfam; PF03464; eRF1\_2; 1.  
DR Pfam; PF03465; eRF1\_3; 1.  
DR TIGRfams; TIGR00111; pefA; 1.  
KW Cell division; Complete proteome.  
SQ SEQUENCE 360 AA; 40814 MW; 2A000CB4B3CE469 CRC64;

Query Match 60.7%; Score 34; DB 17; Length 360;  
Best Local Similarity 45.5%; Pred. No. 1e+02;  
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMHY 11  
Db 98 EDLVPNGSHT 108

## RESULT 22

Q9KBAL  
ID Q9KBAL PRELIMINARY; PRT; 425 AA.  
AC Q9KBAL  
DT 01-OCT-2000 (TReMBLrel. 15, Created)  
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)  
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)  
DE Hypothetical protein BH2027.  
GN BH2027.  
OS Bacillus halodurans.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;  
OC Bacillaceae; Bacillus.  
OX NCBI\_TaxID=86665;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C-125 / JCM 9153;  
RX MEDLINE=20512582; PubMed=11058132;

QY 1 EEVVPXGMHY 10  
Db 218 ENIVPTGKHH 227

## RESULT 24

Q9Y0Y6  
ID Q9Y0Y6 PRELIMINARY; PRT; 637 AA.  
AC Q9Y0Y6

RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,  
RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,  
RA Horikoshi K.;  
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus  
halodurans and genomic sequence comparison with Bacillus subtilis.";  
RL Nucleic Acids Res. 28:4317-4331(2000).  
DR EMBL; AP001514; BAB05745.1;  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 425 AA; 47639 MW; 914824B1FB65E3DE CRC64;

Query Match 60.7%; Score 34; DB 16; Length 425;  
Best Local Similarity 60.0%; Pred. No. 1.2e+02;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMHY 10  
Db 403 EELLIEGMHY 412

## RESULT 23

O16912  
ID O16912 PRELIMINARY; PRT; 495 AA.  
AC O16912  
DT 01-JAN-1998 (TReMBLrel. 05, Created)  
DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)  
DE 01-DEC-2001 (TReMBLrel. 19, Last annotation update)  
GN F10D2.3 protein.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoides;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BRISTOL N2;  
RX MEDLINE=94150718; PubMed=7906398;  
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,  
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,  
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,  
RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,  
RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,  
RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,  
RA Parsons N., Percy C., Rifkin L., Roopra A., Saunders D., Showkeen R.,  
RA Smaldon N., Smith A., Sonhammer E., Staden R., Sulston J.,  
RA Thiermy-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,  
RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohlman P.;  
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
elegans.";  
RL Nature 368:32-38(1994).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BRISTOL N2;  
RX GRAVES T., Wohlmann P., Gillam B.;  
RA "The sequence of C. elegans cosmid F10D2.";  
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF022972; AAC48234.1;  
DR InterPro; IPR004151; Sre.  
DR Pfam; PF03125; Sre; 1.  
SQ SEQUENCE 495 AA; 58190 MW; 0C61139C138EEB4C CRC64;

Query Match 60.7%; Score 34; DB 5; Length 495;  
Best Local Similarity 50.0%; Pred. No. 1.4e+02;  
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMHY 10  
Db 218 ENIVPTGKHH 227



Query Match 60.7%; Score 34; DB 5; Length 637;  
 Best Local Similarity 66.7%; Pred. No. 1.9e+02;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 VVPXGMHYS 11  
 ||| |||

Db 263 VVPDAVHYS 271  
 ||| |||

RESULT 25

Q9E1X6 PRELIMINARY; PRT; 678 AA.

AC Q9E1X6;  
 DT 01-MAR-2001 (TREMBlrel. 16, Created)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE Hypothetical 75.9 kDa protein.  
 OS Cercopitheine herpesvirus 7.  
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
 OC Alphaherpesvirinae; Varicellovirus.  
 OX NCBI\_TaxID=35245;  
 RN [1]

SEQUENCE FROM N.A.  
 RP Gray W.L., Starnes H.B., White M.W., Ashburn C.V., Mahalingam R.;  
 RA "Complete Sequence of the Simian Varicella Virus Genome."  
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF275348; AAG27217.1; -  
 KW Hypothetical protein.  
 SQ SEQUENCE 678 AA; 75850 MW; A17B09E30512FE3C CRC64;

Query Match 60.7%; Score 34; DB 12; Length 678;  
 Best Local Similarity 50.0%; Pred. No. 2e+02;  
 Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMHY 10  
 ||| |||

Db 147 EEIIPKTRY 156

Search completed: June 10, 2003, 13:46:33  
 Job time : 26.7857 secs

DT 01-NOV-1999 (TREMBlrel. 12, Created)  
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE BCDNA:LD28657 protein.  
 GN BCDNA:LD28657 OR CG1098.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 SEQUENCE FROM N.A.  
 RP STRAIN-BERKELEY;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champagne M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.D.G.,  
 RA April J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke S., Mays A.D., Dew I., Dietz S.M.,  
 RA de Pablos B., Delcher A., Deng Z., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster."  
 RL Science 287:2185-2195(2000).  
 RN [2]  
 SEQUENCE FROM N.A.  
 RP Rubin G.M., Wan K.H., Harvey D., Lewis S.E., Brokstein P., Tsang G.,  
 RA Agbayani A., Arcaina T.T., Baxter E., Blazey R.G., Butenhoff C.,  
 RA Champe M., Chavez C., Chew M., Doyle C.M., Farfan D.E., Frise E.,  
 RA Galle R., George R.A., Harris N.L., Hoskins R.A., Evans-Holm M.,  
 RA Houston K.A., Hummasti S.R., Kim E., Li P., Moshrefi M., Pacleb J.M.,  
 RA Park S., Sequeira A., Sethi H., Snir E., Swirskas R.R., Weinburg T.,  
 RA Celniker S.E.;  
 RT "Full length Drosophila melanogaster cDNA sequence."  
 RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AE003601; AAF51961.1; -  
 DR EMBL: AF145690; AAD38665.1; -  
 DR FlyBase: FBgn0027497; BCDNA:LD28657.  
 DR InterPro: IPR000719; Euk pk kinase.  
 DR Pfam: PF00069; pk kinase; 1  
 DR ProDom: PD0000001; Euk\_pk kinase; 1  
 DR PROSITE: PS50011; PROTEIN\_KINASE\_DOM; 1.  
 KW ATP-binding; Transferase.  
 SQ SEQUENCE 637 AA; 70507 MW; ABBB262CAFA44D20 CRC64;



us-09-909-164-12.rag

Wed Jun 11 15:42:23 2003

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2003, 13:24:45 ; Search time 31.3571 Seconds  
(without alignments)  
46.744 Million cell updates/sec

Title: US-09-909-164-12  
Perfect score: 54  
Sequence: 1 EEVVPXGMDYS 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues  
Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 75 summaries

Database :			
A_Geneseq_101002.*			
1:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*		
2:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*		
3:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*		
4:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*		
5:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*		
6:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*		
7:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*		
8:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*		
9:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*		
10:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*		
11:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*		
12:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*		
13:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*		
14:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*		
15:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*		
16:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*		
17:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*		
18:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*		
19:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*		
20:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*		
21:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*		
22:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*		
23:	/SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*		

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	52	96.3	11	23	ABB80524
2	52	96.3	11	23	ABB80528
3	52	96.3	11	23	ABB80529
4	52	96.3	11	23	ABB80561
5	52	96.3	11	23	ABB80562
6	47	87.0	11	23	ABB80538
7	47	87.0	11	23	ABB80542
8	47	87.0	11	23	ABB80543
9	46	85.2	11	23	ABB80521
10	46	85.2	11	23	ABB80522

RESULT 1  
ABB80524  
ID ABB80524 standard; peptide; 11 AA.

11	46	85.2	11	23	ABB80525	Hepatitis C virus
12	46	85.2	11	23	ABB80526	Hepatitis C virus
13	46	85.2	11	23	ABB80547	Hepatitis C virus
14	46	85.2	11	23	ABB80548	Hepatitis C virus
15	46	85.2	11	23	ABB80551	Hepatitis C virus
16	46	85.2	11	23	ABB80556	Hepatitis C virus
17	46	85.2	11	23	ABB80557	Hepatitis C virus
18	46	85.2	11	23	ABB80559	Hepatitis C virus
19	46	85.2	11	23	ABB80563	Hepatitis C virus
20	46	85.2	11	23	ABB80564	Hepatitis C virus
21	46	85.2	11	23	ABB80565	Hepatitis C virus
22	46	85.2	11	23	ABB80566	Hepatitis C virus
23	46	85.2	11	23	ABB80567	Hepatitis C virus
24	46	85.2	11	23	ABB80568	Hepatitis C virus
25	45	83.3	11	23	ABB80527	Hepatitis C virus
26	45	83.3	11	23	ABB80558	Hepatitis C virus
27	45	83.3	11	23	ABB80560	Hepatitis C virus
28	45	83.3	11	23	ABB80560	Hepatitis C virus
29	44	81.5	11	23	ABB80533	Hepatitis C virus
30	44	81.5	11	23	ABB80534	Hepatitis C virus
31	41	75.9	11	23	ABB80535	Hepatitis C virus
32	41	75.9	11	23	ABB80536	Hepatitis C virus
33	41	75.9	11	23	ABB80539	Hepatitis C virus
34	41	75.9	11	23	ABB80540	Hepatitis C virus
35	40	74.1	11	23	ABB80537	Hepatitis C virus
36	40	74.1	11	23	ABB80541	Hepatitis C virus
37	40	74.1	11	23	ABB80544	Hepatitis C virus
38	40	74.1	11	23	ABB80545	Hepatitis C virus
39	40	74.1	11	23	ABB80549	Hepatitis C virus
40	40	74.1	11	23	ABB80552	Hepatitis C virus
41	40	74.1	11	23	ABB80553	Hepatitis C virus
42	39	72.2	11	23	ABB80546	Hepatitis C virus
43	39	72.2	11	23	ABB80550	Hepatitis C virus
44	39	72.2	11	23	ABB80554	Hepatitis C virus
45	39	72.2	11	23	ABB80555	Hepatitis C virus
46	38	70.4	11	23	ABB80530	Hepatitis C virus
47	37	68.5	11	23	ABB80531	Hepatitis C virus
48	37	68.5	11	23	ABB80532	Hepatitis C virus
49	36	66.7	20	20	AAU76810	Hepatitis C virus
50	36	66.7	341	22	AAE00907	Citrus partial suc
51	36	66.7	348	22	AAE00908	Citrus partial suc
52	36	66.7	460	23	AB91579	Herbicideally activ
53	36	66.7	1022	22	ABG03621	Novel human diagno
54	36	66.7	1022	22	ABG05826	Novel human diagno
55	36	66.7	1022	22	ABG08173	Novel human diagno
56	36	66.7	1068	13	AA20198	Sucrose phosphate
57	36	66.7	1068	13	AA27931	SPS protein. Zea
58	36	66.7	1068	18	AAW38266	Maize sucrose phos
59	36	66.7	1068	18	AAW0989	Sucrose phosphate
60	36	66.7	1071	15	AAW0796	Rice sucrose phosph
61	36	66.7	1083	23	AB92875	Herbicideally activ
62	35	64.8	2778	22	AB58683	Drosophila melanog
63	34.5	63.9	748	22	AAW40225	Human polypeptide
64	34	63.0	140	16	AAW4232	MAB L243 VH region
65	34	63.0	140	16	AAW4235	Humanized antibody
66	34	63.0	140	16	AAW4265	CDR-grafted L243-g
67	34	63.0	140	16	AAW4257	MHC-II MAB L243 he
68	34	63.0	143	22	ABW0256	Drosophila melanog
69	34	63.0	150	21	AAW28379	Arabidopsis thalis
70	34	63.0	150	22	AAW87623	Bovine mammary tis
71	34	63.0	222	22	AAU03629	Group B Streptococ
72	34	63.0	222	23	ABP26468	Streptococcus poly
73	34	63.0	565	23	ABW3631	Lactococcus lactis
74	34	63.0	587	21	AAW4287	Neisseria meningit
75	34	63.0	1062	22	AAU33615	Pseudomonas aerugi

ALIGNMENTS

XX AC ABB80524;  
 XX DT 08-OCT-2002 (first entry)  
 XX DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #4.  
 XX KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
 XX OS virucide.  
 XX OS Synthetic.  
 XX FH Key Location/Qualifiers  
 XX FT Modified-site 1 /note= "N-terminal acetyl"  
 XX FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with  
 XX FT residue 7"  
 XX FT Misc-difference 9  
 XX FT Modified-site 11 /note= "D-form residue"  
 XX FT Modified-site 11 /note= "C-terminal amide"  
 XX PN WO200208251-A2.  
 XX PD 31-JAN-2002.  
 XX PF 19-JUL-2001; 2001WO-US23169.  
 XX PR 21-JUL-2000; 2000US-220101P.  
 XX PA (CORV-) CORVAS INT INC.  
 XX PI Lim-wilby M, Levy OE, Brunck TK;  
 XX DR WPI; 2002-361643/39.  
 XX PT Novel peptide compound having hepatitis C virus protease inhibitory  
 XX PT activity useful for treating disorders associated with hepatitis C  
 XX PT virus protease  
 XX PS Claim 17; Page 64; 69pp; English.  
 XX CC The sequence represents a peptide compound of the invention having  
 XX CC Hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
 XX CC invention are alpha-ketoamide peptide analogues. The peptides have  
 XX CC virucide activity, and are useful for treating and in the manufacture of  
 XX CC a medicament to treat disorders associated with HCV protease. A  
 XX CC pharmaceutical composition comprising the peptide as an active ingredient  
 XX CC is useful for treating disorders associated with hepatitis C virus.  
 XX SQ Sequence 11 AA;  
 Query Match 96.3%; Score 52; DB 23; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 0.0015;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 EEVVPXGMDYS 11  
 Db 1 EEVVPXGMDYS 11  
 RESULT 2  
 ABB80528  
 ID ABB80528 standard; peptide; 11 AA.  
 XX AC ABB80528;  
 XX DT 08-OCT-2002 (first entry)  
 XX DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #8.  
 XX KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;

KW virucide.  
 XX Synthetic.  
 XX FH Key Location/Qualifiers  
 XX FT Modified-site 1 /note= "N-terminal acetyl"  
 XX FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with  
 XX FT residue 7"  
 XX FT Misc-difference 8  
 XX FT Modified-site 11 /note= "D-form residue"  
 XX FT Modified-site 11 /note= "C-terminal amide"  
 XX PN WO200208251-A2.  
 XX PD 31-JAN-2002.  
 XX PF 19-JUL-2001; 2001WO-US23169.  
 XX PR 21-JUL-2000; 2000US-220101P.  
 XX PA (CORV-) CORVAS INT INC.  
 XX PI Lim-wilby M, Levy OE, Brunck TK;  
 XX DR WPI; 2002-361643/39.  
 XX PT Novel peptide compound having hepatitis C virus protease inhibitory  
 XX PT activity useful for treating disorders associated with hepatitis C  
 XX PT virus protease  
 XX PS Claim 17; Page 64; 69pp; English.  
 XX CC The sequence represents a peptide compound of the invention having  
 XX CC Hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
 XX CC invention are alpha-ketoamide peptide analogues. The peptides have  
 XX CC virucide activity, and are useful for treating and in the manufacture of  
 XX CC a medicament to treat disorders associated with HCV protease. A  
 XX CC pharmaceutical composition comprising the peptide as an active ingredient  
 XX CC is useful for treating disorders associated with hepatitis C virus.  
 XX SQ Sequence 11 AA;  
 Query Match 96.3%; Score 52; DB 23; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 0.0015;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 EEVVPXGMDYS 11  
 Db 1 EEVVPXGMDYS 11  
 RESULT 3  
 ABB80529  
 ID ABB80529 standard; peptide; 11 AA.  
 XX AC ABB80529;  
 XX DT 08-OCT-2002 (first entry)  
 XX DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #9.  
 XX KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
 XX KW virucide.  
 XX OS Synthetic.  
 XX FH Key Location/Qualifiers  
 XX FT Modified-site 1 /note= "N-terminal acetyl"  
 XX FT Modified-site 6

FT FT /note= "Norvalyl carbonyl forming keto-amide linkage with  
FT FT residue 7"  
FT FT  
FT FT Misc-difference 8 /note= "D-form residue"  
FT FT  
FT FT Misc-difference 9 /note= "D-form residue"  
FT FT  
FT FT Modified-site 11 /note= "C-terminal amide"  
FT FT  
XX XX WO200208251-A2.  
XX XX  
XX XX 31-JAN-2002.  
XX XX  
XX XX 19-JUL-2001; 2001WO-US23169.  
XX XX  
XX XX 21-JUL-2000; 2000US-220101P.  
XX XX  
XX XX (CORV-) CORVAS INT INC.  
XX XX  
XX XX Lim-wilby M, Levy OE, Brunck TK;  
XX XX  
XX XX WPI; 2002-361643/39.  
XX XX  
XX XX Novel peptide compound having hepatitis C virus protease inhibitory  
XX XX activity useful for treating disorders associated with hepatitis C  
XX XX virus protease  
XX XX  
XX XX Claim 17; Page 64; 69pp; English.  
XX XX  
XX XX The sequence represents a peptide compound of the invention having  
XX XX hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
XX XX invention are alpha-ketoamide peptide analogues. The peptides have  
XX XX virucide activity, and are useful for treating and in the manufacture of  
XX XX a medicament to treat disorders associated with HCV protease. A  
XX XX pharmaceutical composition comprising the peptide as an active ingredient  
XX XX is useful for treating disorders associated with hepatitis C virus.  
XX XX  
SQ Sequence 11 AA;  
Query Match 96.3%; Score 52; DB 23; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.0015;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 EEVVPXGMDYS 11  
Db 1 EEVVPXGMDYS 11  
RESULT 4  
ID ABB80561  
XX ABB80561 standard; peptide; 11 AA.  
XX AC ABB80561;  
XX XX  
XX XX 08-OCT-2002 (first entry)  
XX XX  
XX XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #41.  
XX XX  
XX XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
XX XX virucide.  
XX XX  
XX XX Synthetic.  
XX XX  
XX XX Key Location/Qualifiers  
XX XX Modified-site 1 /note= "N-terminal acetyl"  
XX XX  
XX XX Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with  
XX XX residue 7"  
XX XX  
XX XX Misc-difference 8 /note= "D-form residue"  
XX XX  
XX XX Modified-site 8 /note= "Oxymethionine"  
XX XX

FT FT Modified-site 11 /note= "C-terminal amide"  
FT FT  
XX XX WO200208251-A2.  
XX XX  
XX XX 31-JAN-2002.  
XX XX  
XX XX 19-JUL-2001; 2001WO-US23169.  
XX XX  
XX XX 21-JUL-2000; 2000US-220101P.  
XX XX  
XX XX (CORV-) CORVAS INT INC.  
XX XX  
XX XX Lim-wilby M, Levy OE, Brunck TK;  
XX XX  
XX XX WPI; 2002-361643/39.  
XX XX  
XX XX Novel peptide compound having hepatitis C virus protease inhibitory  
XX XX activity useful for treating disorders associated with hepatitis C  
XX XX virus protease  
XX XX  
XX XX Claim 17; Page 65; 69pp; English.  
XX XX  
XX XX The sequence represents a peptide compound of the invention having  
XX XX hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
XX XX invention are alpha-ketoamide peptide analogues. The peptides have  
XX XX virucide activity, and are useful for treating and in the manufacture of  
XX XX a medicament to treat disorders associated with HCV protease. A  
XX XX pharmaceutical composition comprising the peptide as an active ingredient  
XX XX is useful for treating disorders associated with hepatitis C virus.  
XX XX  
SQ Sequence 11 AA;  
Query Match 96.3%; Score 52; DB 23; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.0015;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 EEVVPXGMDYS 11  
Db 1 EEVVPXGMDYS 11  
RESULT 5  
ID ABB80562  
XX ABB80562 standard; peptide; 11 AA.  
XX AC ABB80562;  
XX XX  
XX XX 08-OCT-2002 (first entry)  
XX XX  
XX XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #42.  
XX XX  
XX XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
XX XX virucide.  
XX XX  
XX XX Synthetic.  
XX XX  
XX XX Key Location/Qualifiers  
XX XX Modified-site 1 /note= "N-terminal acetyl"  
XX XX  
XX XX Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with  
XX XX residue 7"  
XX XX  
XX XX Misc-difference 8 /note= "D-form residue"  
XX XX  
XX XX Modified-site 8 /note= "Oxymethionine"  
XX XX  
XX XX Misc-difference 9 /note= "D-form residue"  
XX XX  
XX XX Modified-site 11 /note= "C-terminal amide"  
XX XX  
XX XX WO200208251-A2.  
XX XX

```

XX PD 31-JAN-2002.
XX PF 19-JUL-2001; 2001WO-US23169.
XX PR 21-JUL-2000; 2000US-220101P.
XX PA (CORV-) CORVAS INT INC.
XX PI Lim-wilby M, Levy OE, Brunck TK;
XX DR WPI; 2002-361643/39.
XX CC Novel peptide compound having hepatitis C virus protease inhibitory
XX PT activity useful for treating disorders associated with hepatitis C
XX PT virus protease
XX PS Claim 17; Page 65; 69pp; English.
XX CC The sequence represents a peptide compound of the invention having
XX CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
XX CC invention are alpha-ketoamide peptide analogues. The peptides have
XX CC virucide activity, and are useful for treating and in the manufacture of
XX CC a medicament to treat disorders associated with HCV protease. A
XX CC pharmaceutical composition comprising the peptide as an active ingredient
XX CC is useful for treating disorders associated with hepatitis C virus.
XX SQ Sequence 11 AA;
XX Query Match 96.3%; Score 52; DB 23; Length 11;
XX Best Local Similarity 100.0%; Pred. No. 0.0015;
XX Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX QY 1 EEVVPXGMDYS 11
XX DB 1 EEVVPXGMDYS 11
XX RESULT 6
XX ID ABB80538 standard; peptide; 11 AA.
XX AC ABB80538;
XX DT 08-OCT-2002 (first entry)
XX DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #18.
XX KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
XX KW virucide.
XX OS Synthetic.
XX FH Key Location/Qualifiers
XX FT Modified-site 1 /note= "N-terminal acetyl"
XX FT Modified-site 6 /note= "N-terminal acetyl"
XX FT Misc-difference 9 /note= "Norvalyl carbonyl forming keto-amide linkage with
XX FT residue 7"
XX FT Modified-site 11 /note= "D-form residue"
XX FT Modified-site 11 /note= "C-terminal amide"
XX PN WO200208251-A2.
XX XX 31-JAN-2002.
XX PD 19-JUL-2001; 2001WO-US23169.
XX PF 21-JUL-2000; 2000US-220101P.
XX PR (CORV-) CORVAS INT INC.
XX PA

```

```

XX PI Lim-wilby M, Levy OE, Brunck TK;
XX DR WPI; 2002-361643/39.
XX CC Novel peptide compound having hepatitis C virus protease inhibitory
XX PT activity useful for treating disorders associated with hepatitis C
XX PT virus protease
XX PS Claim 17; Page 64; 69pp; English.
XX CC The sequence represents a peptide compound of the invention having
XX CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
XX CC invention are alpha-ketoamide peptide analogues. The peptides have
XX CC virucide activity, and are useful for treating and in the manufacture of
XX CC a medicament to treat disorders associated with HCV protease. A
XX CC pharmaceutical composition comprising the peptide as an active ingredient
XX CC is useful for treating disorders associated with hepatitis C virus.
XX SQ Sequence 11 AA;
XX Query Match 87.0%; Score 47; DB 23; Length 11;
XX Best Local Similarity 90.9%; Pred. No. 0.014;
XX Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX QY 1 EEVVPXGMDYS 11
XX DB 1 EEVVPXGMDYS 11
XX RESULT 7
XX ID ABB80542 standard; peptide; 11 AA.
XX AC ABB80542;
XX DT 08-OCT-2002 (first entry)
XX DE Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #22.
XX KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
XX KW virucide.
XX OS Synthetic.
XX FH Key Location/Qualifiers
XX FT Modified-site 1 /note= "N-terminal acetyl"
XX FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with
XX FT residue 7"
XX FT Misc-difference 8 /note= "D-form residue"
XX FT Modified-site 11 /note= "C-terminal amide"
XX PN WO200208251-A2.
XX XX 31-JAN-2002.
XX PD 19-JUL-2001; 2001WO-US23169.
XX PF 21-JUL-2000; 2000US-220101P.
XX PR (CORV-) CORVAS INT INC.
XX PA Lim-wilby M, Levy OE, Brunck TK;
XX DR WPI; 2002-361643/39.
XX CC Novel peptide compound having hepatitis C virus protease inhibitory
XX PT activity useful for treating disorders associated with hepatitis C
XX PT virus protease
XX PA

```

CC	virucide activity, and are useful for treating and in the manufacture of
CC	a medicament to treat disorders associated with HCV protease. A
CC	pharmaceutical composition comprising the peptide as an active ingredient
CC	is useful for treating disorders associated with hepatitis C virus.
XX	
XX	
SQ	Sequence 11 AA;
	Query Match 87.0%; Score 47; DB 23; Length 11;
	Best Local Similarity 90.9%; Pred. No. 0.014;
	Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY	1 EEVVPXGMDYS 11
DB	1 EEVVPXGQDYS 11
RESULT 9	
ABB80521	standard; peptide; 11 AA.
XX	
AC	ABB80521;
XX	
DT	08-OCT-2002 (first entry)
XX	
DE	Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #1.
XX	
KW	Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;
KW	virucide.
XX	
OS	Synthetic.
XX	
FH	Key Location/Qualifiers
FT	Modified-site 1
FT	/note= "N-terminal acetyl"
FT	Modified-site 6
FT	/note= "Norvalyl carbonyl forming keto-amide linkage with
FT	residue 7"
FT	Modified-site 11
FT	/note= "C-terminal amide"
XX	
PN	WO200208251-A2.
XX	
PD	31-JAN-2002.
XX	
PF	19-JUL-2001; 2001WO-US23169.
XX	
PR	21-JUL-2000; 2000US-220101P.
XX	
PA	(CORV-) CORVAS INT INC.
XX	
PI	Llm-wilby M, Levy OE, Brunck TK;
XX	
DR	WPI; 2002-361643/39.
XX	
PT	Novel peptide compound having hepatitis C virus protease inhibitory
PT	activity useful for treating disorders associated with hepatitis C
PT	virus protease
XX	
PS	Claim 17; Page 64; 69pp; English.
XX	
CC	The sequence represents a peptide compound of the invention having
CC	hepatitis C virus (HCV) protease inhibitory activity. The peptides of the
CC	invention are alpha-ketoamide peptide analogues. The peptides have
CC	virucide activity, and are useful for treating and in the manufacture of
CC	a medicament to treat disorders associated with HCV protease. A
CC	pharmaceutical composition comprising the peptide as an active ingredient
CC	is useful for treating disorders associated with hepatitis C virus.
XX	
SQ	Sequence 11 AA;
	Query Match 85.2%; Score 46; DB 23; Length 11;
	Best Local Similarity 90.9%; Pred. No. 0.022;
	Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

RESULT 11  
ABB80525





PN WO200208251-A2.  
XX 31-JAN-2002.  
PD 19-JUL-2001; 2001WO-US23169.  
XX 21-JUL-2000; 2000US-220101P.  
PF 21-JUL-2000; 2000US-220101P.  
PR (CORV-) CORVAS INT INC.  
XX Lim-wilby M, Levy OE, Brunck TK;  
PI WPI; 2002-361643/39.  
XX Novel peptide compound having hepatitis C virus protease inhibitory  
XX activity useful for treating disorders associated with hepatitis C  
XX virus protease -  
XX Claim 17; Page 65; 69pp; English.  
XX The sequence represents a peptide compound of the invention having  
XX hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
XX invention are alpha-ketoamide peptide analogues. The peptides have  
XX virucide activity, and are useful for treating and in the manufacture of  
XX a medicament to treat disorders associated with HCV protease. A  
XX pharmaceutical composition comprising the peptide as an active ingredient  
XX is useful for treating disorders associated with hepatitis C virus.  
XX Sequence 11 AA;  
XX Query Match 85.2%; Score 46; DB 23; Length 11;  
XX Best Local Similarity 90.9%; Pred. No. 0.022;  
XX Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 EEVVPXGMDYS 11  
DB 1 EEVVPXGMDYS 11  
RESULT 15  
ABB80551  
ID ABB80551 standard; peptide; 11 AA.  
XX ABB80551;  
XX 08-OCT-2002 (first entry)  
XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #31.  
XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
XX virucide.  
XX Synthetic.  
XX Key Location/Qualifiers  
XX Modified-site 1 /note= "N-terminal acetyl"  
XX Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with  
XX Misc-difference 9 residue 7"  
XX Modified-site 11 /note= "D-form residue"  
XX Modified-site 11 /note= "C-terminal amide"  
XX WO200208251-A2.  
XX 31-JAN-2002.  
XX 19-JUL-2001; 2001WO-US23169.  
XX 21-JUL-2000; 2000US-220101P.  
XX (CORV-) CORVAS INT INC.  
XX Lim-wilby M, Levy OE, Brunck TK;  
XX WPI; 2002-361643/39.  
XX Novel peptide compound having hepatitis C virus protease inhibitory  
XX activity useful for treating disorders associated with hepatitis C  
XX virus protease -  
XX Claim 17; Page 65; 69pp; English.  
XX The sequence represents a peptide compound of the invention having  
XX hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
XX invention are alpha-ketoamide peptide analogues. The peptides have  
XX virucide activity, and are useful for treating and in the manufacture of  
XX a medicament to treat disorders associated with HCV protease. A  
XX pharmaceutical composition comprising the peptide as an active ingredient  
XX is useful for treating disorders associated with hepatitis C virus.  
XX Sequence 11 AA;  
XX Query Match 85.2%; Score 46; DB 23; Length 11;  
XX Best Local Similarity 90.9%; Pred. No. 0.022;  
XX Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 EEVVPXGMDYS 11  
DB 1 EEVVPXGMDYS 11  
RESULT 16  
ABB80556  
ID ABB80556 standard; peptide; 11 AA.  
XX ABB80556;  
XX 08-OCT-2002 (first entry)  
XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #36.  
XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
XX virucide.  
XX Synthetic.  
XX Key Location/Qualifiers  
XX Modified-site 1 /note= "N-terminal acetyl"  
XX Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with  
XX Misc-difference 8 residue 7"  
XX Modified-site 11 /note= "D-form residue"  
XX Modified-site 11 /note= "C-terminal amide"  
XX WO200208251-A2.  
XX 31-JAN-2002.  
XX 19-JUL-2001; 2001WO-US23169.  
XX 21-JUL-2000; 2000US-220101P.  
XX (CORV-) CORVAS INT INC.  
XX Lim-wilby M, Levy OE, Brunck TK;  
XX WPI; 2002-361643/39.  
XX Novel peptide compound having hepatitis C virus protease inhibitory  
XX activity useful for treating disorders associated with hepatitis C  
XX virus protease -

PA (CORV-) CORVAS INT INC.  
XX Lim-wilby M, Levy OE, Brunck TK;  
XX WPI; 2002-361643/39.  
XX Novel peptide compound having hepatitis C virus protease inhibitory  
XX activity useful for treating disorders associated with hepatitis C  
XX virus protease -  
XX Claim 17; Page 65; 69pp; English.  
XX The sequence represents a peptide compound of the invention having  
XX hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
XX invention are alpha-ketoamide peptide analogues. The peptides have  
XX virucide activity, and are useful for treating and in the manufacture of  
XX a medicament to treat disorders associated with HCV protease. A  
XX pharmaceutical composition comprising the peptide as an active ingredient  
XX is useful for treating disorders associated with hepatitis C virus.  
XX Sequence 11 AA;  
XX Query Match 85.2%; Score 46; DB 23; Length 11;  
XX Best Local Similarity 90.9%; Pred. No. 0.022;  
XX Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 EEVVPXGMDYS 11  
DB 1 EEVVPXGMDYS 11  
RESULT 16  
ABB80556  
ID ABB80556 standard; peptide; 11 AA.  
XX ABB80556;  
XX 08-OCT-2002 (first entry)  
XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #36.  
XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
XX virucide.  
XX Synthetic.  
XX Key Location/Qualifiers  
XX Modified-site 1 /note= "N-terminal acetyl"  
XX Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with  
XX Misc-difference 8 residue 7"  
XX Modified-site 11 /note= "D-form residue"  
XX Modified-site 11 /note= "C-terminal amide"  
XX WO200208251-A2.  
XX 31-JAN-2002.  
XX 19-JUL-2001; 2001WO-US23169.  
XX 21-JUL-2000; 2000US-220101P.  
XX (CORV-) CORVAS INT INC.  
XX Lim-wilby M, Levy OE, Brunck TK;  
XX WPI; 2002-361643/39.  
XX Novel peptide compound having hepatitis C virus protease inhibitory  
XX activity useful for treating disorders associated with hepatitis C  
XX virus protease -

PT virus protease -

PS Claim 17; Page 65; 69pp; English.

XX The sequence represents a peptide compound of the invention having  
XX hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
XX invention are alpha-ketoamide peptide analogues. The peptides have  
XX virucide activity, and are useful for treating and in the manufacture of  
XX a medicament to treat disorders associated with HCV protease. A  
XX pharmaceutical composition comprising the peptide as an active ingredient  
XX is useful for treating disorders associated with hepatitis C virus.

SQ Sequence 11 AA;

Query Match 85.2%; Score 46; DB 23; Length 11;

Best Local Similarity 90.9%; Pred. No. 0.022;

Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXGMDYS 11

DB 1 EEVVPXGSDYS 11

RESULT 17

ABB80557

ID ABB80557 standard; peptide; 11 AA.

AC ABB80557;

DT 08-OCT-2002 (first entry)

XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #37.

XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
XX virucide.

OS Synthetic.

FH Key Location/Qualifiers

FT Modified-site 1 /note= "N-terminal acetyl"

FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with  
FT residue 7"

FT Misc-difference 8 /note= "D-form residue"

FT Misc-difference 9 /note= "D-form residue"

FT Modified-site 11 /note= "C-terminal amide"

XX WO200208251-A2.

XX 31-JAN-2002.

XX 19-JUL-2001; 2001WO-US23169.

XX 21-JUL-2000; 2000US-220101P.

XX (CORV-) CORVAS INT INC.

XX Lim-wilby M, Levy OE, Brunck TK;

XX WPI; 2002-361643/39.

XX Novel peptide compound having hepatitis C virus protease inhibitory  
XX activity useful for treating disorders associated with hepatitis C  
XX virus protease

PS Claim 17; Page 65; 69pp; English.

XX The sequence represents a peptide compound of the invention having  
XX hepatitis C virus (HCV) protease inhibitory activity. The peptides of the

CC invention are alpha-ketoamide peptide analogues. The peptides have  
CC virucide activity, and are useful for treating and in the manufacture of  
CC a medicament to treat disorders associated with HCV protease. A  
CC pharmaceutical composition comprising the peptide as an active ingredient  
CC is useful for treating disorders associated with hepatitis C virus.

SQ Sequence 11 AA;

Query Match 85.2%; Score 46; DB 23; Length 11;

Best Local Similarity 90.9%; Pred. No. 0.022;

Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXGMDYS 11

DB 1 EEVVPXGSDYS 11

RESULT 18

ABB80559

ID ABB80559 standard; peptide; 11 AA.

AC ABB80559;

DT 08-OCT-2002 (first entry)

XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #39.

XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
XX virucide.

OS Synthetic.

FH Key Location/Qualifiers

FT Modified-site 1 /note= "N-terminal acetyl"

FT Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with  
FT residue 7"

FT Misc-difference 8 /note= "D-form residue"

FT Modified-site 8 /note= "Oxymethionine"

FT Modified-site 11 /note= "C-terminal amide"

XX WO200208251-A2.

XX 31-JAN-2002.

XX 19-JUL-2001; 2001WO-US23169.

XX 21-JUL-2000; 2000US-220101P.

XX (CORV-) CORVAS INT INC.

XX Lim-wilby M, Levy OE, Brunck TK;

XX WPI; 2002-361643/39.

XX Novel peptide compound having hepatitis C virus protease inhibitory  
XX activity useful for treating disorders associated with hepatitis C  
XX virus protease

PS Claim 17; Page 65; 69pp; English.

XX The sequence represents a peptide compound of the invention having  
XX hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
XX invention are alpha-ketoamide peptide analogues. The peptides have  
XX virucide activity, and are useful for treating and in the manufacture of  
XX a medicament to treat disorders associated with HCV protease. A  
XX pharmaceutical composition comprising the peptide as an active ingredient  
XX is useful for treating disorders associated with hepatitis C virus.

SQ Sequence 11 AA;

Query Match 85.2%; Score 46; DB 23; Length 11;  
 Best Local Similarity 90.9%; Pred. No. 0.022;  
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXGMDYS 11  
 |||||  
 Db 1 EEVVPXGMSYS 11

RESULT 19

ABB80563  
 ID ABB80563 standard; peptide; 11 AA.

XX ABB80563;

XX 08-OCT-2002 (first entry)

XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #43.

XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;

XX virucide.

XX Synthetic.

XX OS

XX Key

XX Modified-site 1

XX Location/Qualifiers

XX /note= "N-terminal acetyl"

XX Modified-site 6

XX /note= "Valyl carbonyl forming keto-amide linkage with

XX residue 7"

XX Modified-site 11

XX /note= "C-terminal amide"

XX WO200208251-A2.

XX 31-JAN-2002.

XX 19-JUL-2001; 2001WO-US23169.

XX 21-JUL-2000; 2000US-220101P.

XX (CORV-) CORVAS INT INC.

XX Lim-wilby M, Levy OE, Brunck TK;

XX WPI; 2002-361643/39.

XX Novel peptide compound having hepatitis C virus protease inhibitory

XX activity useful for treating disorders associated with hepatitis C

XX virus protease

XX Claim 17; Page 65; 69pp; English.

XX The sequence represents a peptide compound of the invention having

XX hepatitis C virus (HCV) protease inhibitory activity. The peptides of the

XX invention are alpha-ketoamide peptide analogues. The peptides have

XX virucide activity, and are useful for treating and in the manufacture of

XX a medicament to treat disorders associated with HCV protease. A

XX pharmaceutical composition comprising the peptide as an active ingredient

XX is useful for treating disorders associated with hepatitis C virus.

XX SQ Sequence 11 AA;

Query Match 85.2%; Score 46; DB 23; Length 11;

Best Local Similarity 90.9%; Pred. No. 0.022;

Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 EEVVPXGMDYS 11

|||||

Db 1 EEVVPXGMSYS 11

Query Match

Best Local Similarity

Matches

QY

Db

SQ

Sequence

11 AA;

Query Match

Best Local Similarity

Matches

QY

Db

SQ

RESULT 20

ABB80564

ID ABB80564 standard; peptide; 11 AA.

XX ABB80564;

XX 08-OCT-2002 (first entry)

XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #44.

XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;

XX virucide.

XX Synthetic.

XX Key

XX Modified-site 1

XX Location/Qualifiers

XX /note= "N-terminal acetyl"

XX Modified-site 6

XX /note= "Leucyl carbonyl forming keto-amide linkage with

XX residue 7"

XX Modified-site 11

XX /note= "C-terminal amide"

XX WO200208251-A2.

XX 31-JAN-2002.

XX 19-JUL-2001; 2001WO-US23169.

XX 21-JUL-2000; 2000US-220101P.

XX (CORV-) CORVAS INT INC.

XX Lim-wilby M, Levy OE, Brunck TK;

XX WPI; 2002-361643/39.

XX Novel peptide compound having hepatitis C virus protease inhibitory

XX activity useful for treating disorders associated with hepatitis C

XX virus protease

XX Claim 17; Page 65; 69pp; English.

XX The sequence represents a peptide compound of the invention having

XX hepatitis C virus (HCV) protease inhibitory activity. The peptides of the

XX invention are alpha-ketoamide peptide analogues. The peptides have

XX virucide activity, and are useful for treating and in the manufacture of

XX a medicament to treat disorders associated with HCV protease. A

XX pharmaceutical composition comprising the peptide as an active ingredient

XX is useful for treating disorders associated with hepatitis C virus.

XX SQ Sequence 11 AA;

Query Match

Best Local Similarity

Matches

QY

Db

SQ

Sequence

11 AA;

Query Match

Best Local Similarity

Matches

QY

Db

SQ

Sequence

11 AA;

Query Match

Best Local Similarity

Matches

QY

Db

SQ

Sequence

11 AA;

Query Match

Best Local Similarity

Matches

QY

Db

SQ

Sequence

11 AA;

Query Match

Best Local Similarity

Matches

QY

Db

SQ

Sequence

11 AA;

Query Match

Best Local Similarity

Matches

QY

Db

SQ

Sequence

11 AA;

Query Match

Best Local Similarity

Matches

QY

Db

SQ

Sequence

11 AA;

Query Match

Best Local Similarity

Matches

QY

Db

SQ

Sequence

11 AA;

Query Match

Best Local Similarity

Matches

QY

Db

SQ

Sequence

11 AA;

Query Match

Best Local Similarity

Matches

QY

Db

SQ

Sequence

11 AA;

Query Match

Best Local Similarity

Matches

QY

Db

SQ

Sequence

11 AA;

Query Match

Best Local Similarity

Matches

QY

Db

SQ

Sequence

11 AA;

Query Match

Best Local Similarity

Matches

QY

Db

SQ

Sequence

11 AA;

Query Match

Best Local Similarity

Matches

QY

Db

SQ

Sequence

11 AA;

Query Match

Best Local Similarity

Matches

QY

Db

SQ

Sequence

11 AA;

Query Match

Best Local Similarity

Matches

QY

Db

SQ

Sequence

11 AA;

Query Match

Best Local Similarity

Matches

QY

Db

SQ

Sequence

11 AA;

Query Match

Best Local Similarity

Matches

QY

Db

SQ

Sequence

11 AA;

Query Match

Best Local Similarity

Matches

QY

Db

SQ

Sequence

11 AA;

Query Match

Best Local Similarity

Matches

QY

Db

SQ

Sequence

11 AA;

Query Match

Best Local Similarity

Matches

QY

Db

SQ

Sequence

11 AA;

Query Match

Best Local Similarity

Matches

QY

Db

SQ

Sequence

11 AA;

Query Match

Best Local Similarity

Matches

QY

Db

SQ

Sequence

11 AA;

Query Match

Best Local Similarity

Matches

QY

Db

SQ

Sequence

11 AA;

Query Match

Best Local Similarity

Matches

QY

Db

SQ

Sequence

11 AA;

Query Match

Best Local Similarity

Matches

QY

Db

SQ

Sequence

11 AA;

Query Match

Best Local Similarity

Matches

QY

Db

SQ

Sequence

11 AA;

Query Match

Best Local Similarity

Matches

QY

Db

SQ

Sequence

11 AA;

Query Match

Best Local Similarity

Matches

QY

Db

SQ

Sequence

11 AA;

Query Match

Best Local Similarity

Matches

QY

Db

SQ

Sequence

11 AA;

Query Match

Best Local Similarity

Matches

QY

Db

SQ

Sequence

11 AA;

Query Match

Best Local Similarity

Matches

QY&lt;/

XX KW Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
 XX virucide.  
 XX OS Synthetic.  
 XX FH Key Location/Qualifiers  
 FT Modified-site 1  
 FT /note= "N-terminal acetyl"  
 FT Modified-site 6  
 FT /note= "Norleucyl carbonyl forming keto-amide linkage  
 FT with residue 7"  
 FT Modified-site 11  
 FT /note= "C-terminal amide"  
 FT WO200208251-A2.  
 XX 31-JAN-2002.  
 XX 19-JUL-2001; 2001WO-US23169.  
 XX 21-JUL-2000; 2000US-220101P.  
 XX (CORV-) CORVAS INT INC.  
 XX Lim-wilby M, Levy OE, Brunck TK;  
 XX WPI; 2002-361643/39.  
 XX Novel peptide compound having hepatitis C virus protease inhibitory  
 PT activity useful for treating disorders associated with hepatitis C  
 PT virus protease  
 XX Claim 17; Page 65; 69pp; English.  
 XX The sequence represents a peptide compound of the invention having  
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
 CC invention are alpha-ketoamide peptide analogues. The peptides have  
 CC virucide activity, and are useful for treating and in the manufacture of  
 CC a medicament to treat disorders associated with HCV protease. A  
 CC pharmaceutical composition comprising the peptide as an active ingredient  
 CC is useful for treating disorders associated with hepatitis C virus.  
 XX SQ Sequence 11 AA;  
 Query Match 85.2%; Score 46; DB 23; Length 11;  
 Best Local Similarity 90.9%; Pred. No. 0.022;  
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 EEVVPXGMDYS 11  
 DB 1 EEVVPXGMSYS 11  
 RESULT 22  
 ABB80566  
 ID ABB80566 standard; peptide; 11 AA.  
 XX ABB80566;  
 XX 08-OCT-2002 (first entry)  
 XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #46.  
 XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
 KW virucide.  
 XX Synthetic.  
 XX Key Location/Qualifiers  
 FT Modified-site 1  
 FT /note= "N-terminal acetyl"  
 FT Modified-site 6

FT /note= "2-aminoisobutyl carbonyl residue forming a  
 FT keto-amide linkage with residue 7"  
 FT Modified-site 11  
 FT /note= "C-terminal amide"  
 XX WO200208251-A2.  
 XX 31-JAN-2002.  
 XX 19-JUL-2001; 2001WO-US23169.  
 XX 21-JUL-2000; 2000US-220101P.  
 XX (CORV-) CORVAS INT INC.  
 XX Lim-wilby M, Levy OE, Brunck TK;  
 XX WPI; 2002-361643/39.  
 XX Novel peptide compound having hepatitis C virus protease inhibitory  
 PT activity useful for treating disorders associated with hepatitis C  
 PT virus protease  
 XX Claim 17; Page 65; 69pp; English.  
 XX The sequence represents a peptide compound of the invention having  
 CC hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
 CC invention are alpha-ketoamide peptide analogues. The peptides have  
 CC virucide activity, and are useful for treating and in the manufacture of  
 CC a medicament to treat disorders associated with HCV protease. A  
 CC pharmaceutical composition comprising the peptide as an active ingredient  
 CC is useful for treating disorders associated with hepatitis C virus.  
 XX SQ Sequence 11 AA;  
 Query Match 85.2%; Score 46; DB 23; Length 11;  
 Best Local Similarity 90.9%; Pred. No. 0.022;  
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 EEVVPXGMDYS 11  
 DB 1 EEVVPXGMSYS 11  
 RESULT 23  
 ABB80567  
 ID ABB80567 standard; peptide; 11 AA.  
 XX ABB80567;  
 XX 08-OCT-2002 (first entry)  
 XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #47.  
 XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
 KW virucide.  
 XX Synthetic.  
 XX Key Location/Qualifiers  
 FT Modified-site 1  
 FT /note= "N-terminal acetyl"  
 FT Modified-site 6  
 FT /note= "(s,s)allothreonyl carbonyl residue forming a  
 FT keto-amide linkage with residue 7"  
 FT Modified-site 11  
 FT /note= "C-terminal amide"  
 XX WO200208251-A2.  
 XX 31-JAN-2002.  
 XX 19-JUL-2001; 2001WO-US23169.

XX 21-JUL-2000; 2000US-220101P.  
XX (CORV-) CORVAS INT INC.  
XX  
XX Lim-wilby M, Levy OE, Brunck TK;  
XX WPI; 2002-361643/39.  
XX Novel peptide compound having hepatitis C virus protease inhibitory  
XX activity useful for treating disorders associated with hepatitis C  
XX virus protease  
XX  
XX Claim 17; Page 65; 69pp; English.  
XX  
XX The sequence represents a peptide compound of the invention having  
XX hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
XX invention are alpha-ketoamide peptide analogues. The peptides have  
XX virucide activity, and are useful for treating and in the manufacture of  
XX a medicament to treat disorders associated with HCV protease. A  
XX pharmaceutical composition comprising the peptide as an active ingredient  
XX is useful for treating disorders associated with hepatitis C virus.  
XX  
XX Sequence 11 AA;  
XX  
XX Query Match 85.2%; Score 46; DB 23; Length 11;  
XX Best Local Similarity 90.9%; Pred. No. 0.022; Indels 0; Gaps 0;  
XX Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 EEVVPXGMDYS 11  
DB 1 EEVVPXGMSYS 11  
RESULT 24  
ABB80568  
ID ABB80568 standard; peptide; 11 AA.  
XX AC ABB80568;  
XX  
XX 08-OCT-2002 (first entry)  
XX  
XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #48.  
XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
XX virucide.  
XX Synthetic.  
XX  
XX Key Location/Qualifiers  
XX Modified-site 1 /note= "N-terminal acetyl"  
XX Modified-site 6 /note= "Alpha-propionyl-glycyl-carbonyl residue forming  
XX a keto-amide linkage with residue 7"  
XX Modified-site 11 /note= "C-terminal amide"  
XX  
XX WO200208251-A2.  
XX  
XX 31-JAN-2002.  
XX  
XX 19-JUL-2001; 2001WO-US23169.  
XX  
XX 21-JUL-2000; 2000US-220101P.  
XX (CORV-) CORVAS INT INC.  
XX  
XX Lim-wilby M, Levy OE, Brunck TK;  
XX WPI; 2002-361643/39.  
XX Novel peptide compound having hepatitis C virus protease inhibitory

PT activity useful for treating disorders associated with hepatitis C  
XX virus protease  
XX  
XX Claim 17; Page 65; 69pp; English.  
XX  
XX The sequence represents a peptide compound of the invention having  
XX hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
XX invention are alpha-ketoamide peptide analogues. The peptides have  
XX virucide activity, and are useful for treating and in the manufacture of  
XX a medicament to treat disorders associated with HCV protease. A  
XX pharmaceutical composition comprising the peptide as an active ingredient  
XX is useful for treating disorders associated with hepatitis C virus.  
XX  
XX Sequence 11 AA;  
XX  
XX Query Match 85.2%; Score 46; DB 23; Length 11;  
XX Best Local Similarity 90.9%; Pred. No. 0.022; Indels 0; Gaps 0;  
XX Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 EEVVPXGMDYS 11  
DB 1 EEVVPXGMSYS 11

RESULT 25  
ABB80523  
ID ABB80523 standard; peptide; 11 AA.  
XX AC ABB80523;  
XX  
XX 08-OCT-2002 (first entry)  
XX  
XX Hepatitis C virus NS3/NS4a serine protease inhibitor peptide #3.  
XX Hepatitis C virus; HCV; serine protease; inhibitor; alpha-ketoamide;  
XX virucide.  
XX Synthetic.  
XX  
XX Key Location/Qualifiers  
XX Modified-site 1 /note= "N-terminal acetyl"  
XX Modified-site 6 /note= "Norvalyl carbonyl forming keto-amide linkage with  
XX residue 7"  
XX Misc-difference 9 /note= "D-form residue"  
XX Modified-site 11 /note= "C-terminal amide"  
XX  
XX WO200208251-A2.  
XX  
XX 31-JAN-2002.  
XX  
XX 19-JUL-2001; 2001WO-US23169.  
XX  
XX 21-JUL-2000; 2000US-220101P.  
XX (CORV-) CORVAS INT INC.  
XX  
XX Lim-wilby M, Levy OE, Brunck TK;  
XX WPI; 2002-361643/39.  
XX Novel peptide compound having hepatitis C virus protease inhibitory  
XX activity useful for treating disorders associated with hepatitis C  
XX virus protease  
XX  
XX Claim 17; Page 64; 69pp; English.  
XX  
XX The sequence represents a peptide compound of the invention having  
XX hepatitis C virus (HCV) protease inhibitory activity. The peptides of the  
XX invention are alpha-ketoamide peptide analogues. The peptides have

CC virucide activity, and are useful for treating and in the manufacture of  
CC a medicament to treat disorders associated with HCV protease. A  
CC pharmaceutical composition comprising the peptide as an active ingredient  
CC is useful for treating disorders associated with hepatitis C virus.

XX

SQ Sequence 11 AA;

Query Match 83.3%; Score 45; DB 23; Length 11;

Best Local Similarity 90.9%; Pred. No. 0.034;

Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 EEVVPXGMDYS 11

|||||||

Db 1 EEVVPXGMHYS 11

Search completed: June 10, 2003, 13:39:09

Job time : 32.3571 secs





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

# OM protein - protein search, using sw model

Run on: June 10, 2003, 13:31:45 ; Search time 9.64286 seconds  
(without alignments)  
33.564 Million cell updates/sec

Title: US-09-909-164-12  
Perfect score: 54  
Sequence: 1 EEVVPXGMDYS 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 75 summaries

## Database :

Issued\_Patents\_AA.\*  
1: /cgn2\_6/ptodata/1/1aa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/1aa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/1aa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/1aa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/1aa/PCUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/1aa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	36	66.7	341	US-08-853-948B-4	Sequence 4, Appl
2	36	66.7	348	US-08-853-948B-5	Sequence 5, Appl
3	36	66.7	368	US-09-697-367-24	Sequence 24, Appl
4	36	66.7	1068	US-08-429-054A-11	Sequence 11, Appl
5	36	66.7	1068	US-08-718-777-7	Sequence 7, Appl
6	36	66.7	1068	US-09-051-341-7	Sequence 7, Appl
7	34	63.0	140	US-08-569-147-76	Sequence 76, Appl
8	34	63.0	140	US-08-569-147-82	Sequence 82, Appl
9	33	61.1	59	US-08-963-851-14	Sequence 14, Appl
10	33	61.1	378	US-08-070-165F-8	Sequence 8, Appl
11	33	61.1	378	US-08-885-418-8	Sequence 8, Appl
12	33	61.1	801	US-09-383-630-6	Sequence 6, Appl
13	32	59.3	65	5177197-51	Patent No. 5177197
14	32	59.3	102	US-08-580-988A-23	Sequence 23, Appl
15	32	59.3	152	US-08-460-694-4	Sequence 4, Appl
16	32	59.3	152	US-08-460-744-4	Sequence 4, Appl
17	32	59.3	152	US-07-667-711B-4	Sequence 4, Appl
18	32	59.3	173	US-08-193-977-7	Sequence 7, Appl
19	32	59.3	189	US-08-464-517-21	Sequence 21, Appl
20	32	59.3	189	US-08-246-361A-21	Sequence 21, Appl
21	32	59.3	189	US-08-463-772-21	Sequence 21, Appl
22	32	59.3	189	PCT-US93-05000-21	Sequence 21, Appl
23	32	59.3	231	US-08-926-842B-20	Sequence 20, Appl
24	32	59.3	236	US-08-464-517-22	Sequence 22, Appl
25	32	59.3	236	US-08-246-361A-22	Sequence 22, Appl
26	32	59.3	236	US-08-463-772-22	Sequence 22, Appl
27	32	59.3	236	PCT-US93-05000-22	Sequence 22, Appl

28	32	59.3	240	3	US-08-926-842B-21	Sequence 21, Appl
29	32	59.3	280	2	US-08-464-517-6	Sequence 6, Appl
30	32	59.3	280	2	US-08-463-772-6	Sequence 6, Appl
31	32	59.3	289	2	US-08-246-361A-4	Sequence 4, Appl
32	32	59.3	289	5	PCT-US93-05000-4	Sequence 4, Appl
33	32	59.3	291	5	PCT-US93-05000-6	Sequence 6, Appl
34	32	59.3	292	2	US-08-464-517-23	Sequence 23, Appl
35	32	59.3	292	2	US-08-246-361A-6	Sequence 6, Appl
36	32	59.3	292	2	US-08-246-361A-23	Sequence 23, Appl
37	32	59.3	292	2	US-08-463-772-23	Sequence 23, Appl
38	32	59.3	292	5	PCT-US93-05000-23	Sequence 23, Appl
39	32	59.3	295	1	US-07-947-120-8	Sequence 8, Appl
40	32	59.3	295	1	US-08-472-893A-8	Sequence 8, Appl
41	32	59.3	295	2	US-08-460-694-2	Sequence 2, Appl
42	32	59.3	295	2	US-08-464-517-19	Sequence 19, Appl
43	32	59.3	295	2	US-08-464-517-20	Sequence 20, Appl
44	32	59.3	295	2	US-08-246-361A-19	Sequence 19, Appl
45	32	59.3	295	2	US-08-246-361A-20	Sequence 20, Appl
46	32	59.3	295	3	US-08-463-772-19	Sequence 19, Appl
47	32	59.3	295	3	US-08-463-772-20	Sequence 20, Appl
48	32	59.3	295	3	US-08-460-744-2	Sequence 2, Appl
49	32	59.3	295	3	US-07-667-711B-2	Sequence 2, Appl
50	32	59.3	295	3	US-08-947-492-8	Sequence 8, Appl
51	32	59.3	295	5	PCT-US93-05000-2	Sequence 2, Appl
52	32	59.3	295	5	PCT-US93-05000-19	Sequence 19, Appl
53	32	59.3	295	5	PCT-US93-05000-20	Sequence 20, Appl
54	32	59.3	309	2	US-08-464-517-4	Sequence 4, Appl
55	32	59.3	309	3	US-08-463-772-4	Sequence 4, Appl
56	32	59.3	410	6	5177197-1	Patent No. 5177197
57	32	59.3	493	4	US-09-411-628-10	Sequence 10, Appl
58	32	59.3	618	2	US-08-770-761A-3	Sequence 3, Appl
59	32	59.3	647	2	US-08-770-761A-8	Sequence 8, Appl
60	32	59.3	660	2	US-08-770-761A-2	Sequence 2, Appl
61	32	59.3	662	2	US-08-770-761A-5	Sequence 5, Appl
62	32	59.3	705	2	US-08-770-761A-7	Sequence 7, Appl
63	32	59.3	756	4	US-09-085-199B-9	Sequence 9, Appl
64	32	59.3	819	2	US-08-464-517-7	Sequence 7, Appl
65	32	59.3	819	3	US-08-246-361A-7	Sequence 7, Appl
66	32	59.3	819	3	US-08-463-772-7	Sequence 7, Appl
67	32	59.3	819	5	PCT-US93-05000-7	Sequence 7, Appl
68	32	59.3	914	4	US-09-085-199B-4	Sequence 4, Appl
69	32	59.3	1090	4	US-09-085-199B-5	Sequence 5, Appl
70	32	59.3	1394	6	5177197-30	Patent No. 5177197
71	31	57.4	139	4	US-09-134-001C-5124	Sequence 5124, Ap
72	31	57.4	267	4	US-09-399-913-57	Sequence 57, Appl
73	31	57.4	381	4	US-09-134-001C-3003	Sequence 3003, Ap
74	31	57.4	529	4	US-09-240-639-4	Sequence 4, Appl
75	31	57.4	622	2	US-08-459-146-2	Sequence 2, Appl

## ALIGNMENTS

RESULT 1  
US-08-853-948B-4  
; Sequence 4, Application US/08853948B  
; Patent No. 6210943  
; GENERAL INFORMATION:  
; APPLICANT: AKIHAMA, Toyota  
; TITLE OF INVENTION: SUCROSE PHOSPHATE SYNTHASE FROM CITRUS AND DNA ENCODING  
; TITLE OF INVENTION: THE SAME  
; FILE REFERENCE: 0049-0235-0  
; CURRENT APPLICATION NUMBER: US/08/853,948B  
; CURRENT FILING DATE: 1997-05-09  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 341  
; TYPE: PRT  
; ORGANISM: Citrus unshiu  
; FEATURE:  
; OTHER INFORMATION: Xaa at position 109 is one of Ala, Arg, Asn, Asp,  
; Cys, Gln, Glu, Gly, His, Ile, Leu, Lys, Met, Phe,

OTHER INFORMATION: Pro, Ser, Thr, Trp, Tyr, or Val  
US-08-853-948B-4

Query Match 66.7%; Score 36; DB 4; Length 341;  
Best Local Similarity 66.7%; Pred. No. 14;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMDYS 11  
I:| | | | |  
Db 228 VIPPGMDFS 236

## RESULT 2

US-08-853-948B-5  
; Sequence 5, Application US/08853948B  
; Patent No. 6210943  
; GENERAL INFORMATION:  
; APPLICANT: AKIHAMA, Toyota  
; TITLE OF INVENTION: SUCROSE PHOSPHATE SYNTHASE FROM CITRUS AND DNA ENCODING  
; FILE REFERENCE: THE SAME  
; CURRENT APPLICATION NUMBER: US/08/853,948B  
; CURRENT FILING DATE: 1997-05-09  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 5  
; LENGTH: 348  
; TYPE: PRT  
; ORGANISM: Citrus unshiu  
US-08-853-948B-5

Query Match 66.7%; Score 36; DB 4; Length 348;  
Best Local Similarity 66.7%; Pred. No. 15;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMDYS 11  
I:| | | | |  
Db 234 VIPPGMDFS 242

## RESULT 3

US-09-697-367-24  
; Sequence 24, Application US/09697367  
; Patent No. 6323015  
; GENERAL INFORMATION:  
; APPLICANT: Orozco Jr., Emil M.  
; APPLICANT: Caimi, Perry G.  
; APPLICANT: Weng, Zude  
; APPLICANT: Tarczynski, Mitchell  
; TITLE OF INVENTION: SUCROSE PHOSPHATE SYNTHASE  
; FILE REFERENCE: BB1166 US NA  
; CURRENT APPLICATION NUMBER: US/09/697,367  
; CURRENT FILING DATE: 2000-10-26  
; PRIOR APPLICATION NUMBER: 60/084,529  
; PRIOR FILING DATE: 1998-MAY-07  
; PRIOR APPLICATION NUMBER: PCT/US99/09865  
; PRIOR FILING DATE: 1999-MAY-06  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 24  
; LENGTH: 368  
; TYPE: PRT  
; ORGANISM: Zea mays  
US-09-697-367-24

Query Match 66.7%; Score 36; DB 4; Length 368;  
Best Local Similarity 66.7%; Pred. No. 16;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMDYS 11  
I:| | | | |  
Db 217 VIPPGMDFS 225

## RESULT 4

US-08-429-054A-11  
; Sequence 11, Application US/08429054A  
; Patent No. 5917126  
; GENERAL INFORMATION:  
; APPLICANT: VAN ASSCHE, CHARLES; LANDO, DANIELLE; BRUNEAU,  
; APPLICANT: JEAN; VOELKER, TONI; GERVAIS, MONICA  
; TITLE OF INVENTION: SUCROSE PHOSPHATE SYNTHETASE (SPS),  
; TITLE OF INVENTION: A PREPARATION METHOD AND CDNA THEREFOR, AND USE OF THE  
; TITLE OF INVENTION: CDNA FOR MODIFYING SPS EXPRESSION IN PLANT CELLS  
; NUMBER OF SEQUENCES: 37  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BIERMAN AND MUSERLIAN  
; STREET: 600 THIRD AVENUE  
; CITY: NEW YORK  
; STATE: NEW YORK  
; COUNTRY: USA  
; ZIP: 10016  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: FLOPPY DISK  
; COMPUTER: IBM PC COMPATIBLE  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WORDPERFECT 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/429,054A  
; FILING DATE: 26-APR-1995  
; CLASSIFICATION: 800  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 842,337  
; FILING DATE: 20-March-1992  
; APPLICATION NUMBER: PCT/FR 91/00593  
; FILING DATE: 18-July-1991  
; CLASSIFICATION: 800  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: French 90402094.9  
; FILING DATE: 20-July-1990  
; CLASSIFICATION: 800  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Charles A. Muserlian  
; REGISTRATION NUMBER: 19,683  
; REFERENCE/DOCKET NUMBER: 146.1137  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 661-8000  
; TELEFAX: (212) 661-8002  
; INFORMATION FOR SEQ ID NO: 11:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1068  
; TYPE: Amino acid  
; STRANDEDNESS: Single  
; TOPOLOGY: Unknown  
; MOLECULE TYPE: Peptide  
US-08-429-054A-11

Query Match 66.7%; Score 36; DB 2; Length 1068;  
Best Local Similarity 66.7%; Pred. NO. 53;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMDYS 11  
I:| | | | |  
Db 435 VIPPGMDFS 443

## RESULT 5

US-08-718-777-7  
; Sequence 7, Application US/08718777  
; Patent No. 5981852  
; GENERAL INFORMATION:  
; APPLICANT: Van Assche, C.  
; APPLICANT: Lando, D.  
; APPLICANT: Bruneau, J. M.  
; APPLICANT: Voelker, T.  
; APPLICANT: Gervais, M.

```

; TITLE OF INVENTION: MODIFICATION OF SUCROSE
; TITLE OF INVENTION: PHOSPHATE
; TITLE OF INVENTION: SYNTHASE IN PLANTS
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Law Offices of Barbara Rae-Venter
; STREET: 260 Sheridan Avenue, Suite 440
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/718,777
; FILING DATE: NOT YET ASSIGNED
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/175,471
; FILING DATE: 27-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Barbara Rae-Venter
; REGISTRATION NUMBER: 32,750
; REFERENCE/DOCKET NUMBER: CGNE.072.02US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415)328-4400
; TELEFAX: (415)328-4477
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1068 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-718-777-7

```

```

Query Match 66.7%; Score 36; DB 2; Length 1068;
Best Local Similarity 66.7%; Pred. No. 53;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 3 VVPXGMDYS 11
   1:11111:1
Db 435 VIPPGMDFS 443

```

```

RESULT 6
US-09-051-341-7
; Sequence 7, Application US/09051341
; Patent No. 6124528
; GENERAL INFORMATION:
; APPLICANT: Shewmaker, C. K.
; TITLE OF INVENTION: MODIFICATION OF SOLUBLE SOLIDS USING
; TITLE OF INVENTION: SUCROSE PHOSPHATE SYNTHASE ENCODING SEQUENCE
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rae-Venter Law Group, P.C.
; STREET: 260 Sheridan Avenue, Suite 440
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/051,341
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:

```

```

; APPLICATION NUMBER: PCT/US96/17351
; FILING DATE: 25-OCT-1996
; APPLICATION NUMBER: US 08/549,016
; FILING DATE: 27-OCT-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/372,200
; FILING DATE: 12-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Barbara Rae-Venter, Ph.D.,
; REGISTRATION NUMBER: 32,750
; REFERENCE/DOCKET NUMBER: CGNE.110.02US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415)328-4400
; TELEFAX: (415)328-4477
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1068 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-051-341-7

```

```

Query Match 66.7%; Score 36; DB 3; Length 1068;
Best Local Similarity 66.7%; Pred. No. 53;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 3 VVPXGMDYS 11
   1:11111:1
Db 435 VIPPGMDFS 443

```

```

RESULT 7
US-08-569-147-76
; Sequence 76, Application US/08569147
; Patent No. 6180377
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: HUMANISED ANTIBODIES
; NUMBER OF SEQUENCES: 95
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &
; STREET: No. 6180377ris, LLP
; CITY: Philadelphia
; STATE: PA
; COUNTRY: U.S.A.
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/569,147
; FILING DATE: 25-March-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Trujillo, Doreen Yatko
; REGISTRATION NUMBER: 35,719
; REFERENCE/DOCKET NUMBER: CARP-0047
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 76:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 140 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-569-147-76

```

```

Query Match 63.0%; Score 34; DB 4; Length 140;
Best Local Similarity 75.0%; Pred. No. 13;

```



```
; CITY: Columbus
; STATE: Ohio
; COUNTRY: USA
; ZIP: 43210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/885,418
; FILING DATE:
; CLASSIFICATION: 435
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (614)-293-8093
; TELEFAX: (614)-293-5631
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 378 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-885-418-8

Query Match 61.1%; Score 33; DB 2; Length 378;
Best Local Similarity 85.7%; Pred. No. 64;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 PXGMDYS 11
Db 139 PPGMDYS 145

RESULT 12
US-09-383-630-6
; Sequence 6, Application US/09383630A
; Patent No. 6265632
; GENERAL INFORMATION:
; APPLICANT: Avner Yayon et al.
; TITLE OF INVENTION: ANIMAL MODEL FOR FIBROBLAST GROWTH
; FACTOR RECEPTOR ASSOCIATED
; CHONDRODYSPLASIA
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Mark M. Friedman c/o Anthony Castorina
; STREET: 2001 Jefferson Davis Highway, Suite 207
; CITY: Arlington
; STATE: Virginia
; COUNTRY: United States of America
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 1.44 megabyte, 3.5" microdisk
; COMPUTER: Twinhead* Slimnote-890TX
; OPERATING SYSTEM: MS DOS version 6.2,
; Windows version 3.11
; SOFTWARE: Word for Windows version 2.0 converted
; to an ASCII file
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/383,630A
; FILING DATE: 26-Aug-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: <Unknown>
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Friedman, Mark M.
; REGISTRATION NUMBER: 33,883
; REFERENCE/DOCKET NUMBER: 1402/2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 972-3-562553
; TELEFAX: 972-3-562554
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 6:
```

```
; SEQUENCE CHARACTERISTICS:
; LENGTH: 801
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-383-630-6

Query Match 61.1%; Score 33; DB 4; Length 801;
Best Local Similarity 85.7%; Pred. No. 1.5e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 PXGMDYS 11
Db 566 PPGMDYS 572

RESULT 13
5177197-51
; Patent No. 5177197
; APPLICANT: KANZAKI, TETSUTO; OLOFSSON, ANDERS; MOREN, ANITA;
; WERNSTEDT, CHRISTER; HELLMAN, ULF; MIYAZONO, KOHEI; CLAESSEN-WELSH,
; LENA; HELDIN, CARL-HENRIK
; TITLE OF INVENTION: ISOLATED NUCLEOTIDE SEQUENCE EXPRESSING
; HUMAN TRANSFORMING GROWTH FACTOR-BETAL-BINDING PROTEIN
; NUMBER OF SEQUENCES: 53
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/487,343
; FILING DATE: 27-FEB-1990
; SEQ ID NO: 51:
; LENGTH: 65
; 5177197-51

Query Match 59.3%; Score 32; DB 6; Length 65;
Best Local Similarity 45.5%; Pred. No. 13;
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMDYS 11
Db 52 KEICPGMGYT 62

RESULT 14
US-08-580-988A-23
; Sequence 23, Application US/08580988A
; Patent No. 5856161
; GENERAL INFORMATION:
; APPLICANT: Aggarwal et al.
; TITLE OF INVENTION: Tumor Necrosis Factor
; TITLE OF INVENTION: Receptor-I-Associated Protein Kinase And Methods
; TITLE OF INVENTION: For Its Use
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dr. Benjamin A. Adler
; STREET: 8011 Candle Lane
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 1.44 Mb floppy disk
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh
; SOFTWARE: Microsoft Word for Macintosh
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/580,988A
; FILING DATE: January 3, 1996
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Benjamin Aaron Adler, Ph.D., J.D.
```

REGISTRATION NUMBER: 35,423  
REFERENCE/DOCKET NUMBER: D5721CIP2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 713-777-2321  
TELEFAX: 713-777-6908  
INFORMATION FOR SEQ ID NO: 23:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 102 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
DESCRIPTION: no  
HYPOTHETICAL: no  
ANTI-SENSE: no  
FRAGMENT TYPE: internal  
ORIGINAL SOURCE:  
US-08-580-988A-23

Query Match 59.3%; Score 32; DB 2; Length 102;  
Best Local Similarity 60.0%; Pred. No. 22;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 EEVVPXGMDY 10  
Db 24 EEVFLAMNY 33

RESULT 15  
US-08-460-694-4  
Sequence 4, Application US/08460694  
Patent No. 5858655  
GENERAL INFORMATION:  
APPLICANT: Arnold, Andrew  
TITLE OF INVENTION: Pradl Cyclin and its cdna  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
STREET: 1100 New York Avenue, N.W., Suite 600  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/460,694  
FILING DATE: 02-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: McConathy, Evelyn H.  
REGISTRATION NUMBER: 35,279  
REFERENCE/DOCKET NUMBER: 0609.4070002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-2600  
TELEFAX: 202-371-2540  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 152 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-460-694-4

Query Match 59.3%; Score 32; DB 2; Length 152;  
Best Local Similarity 60.0%; Pred. No. 35;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 EEVVPXGMDY 10

Db 20 EEVFLAMNY 29

RESULT 16  
US-08-460-744-4  
Sequence 4, Application US/08460744  
Patent No. 6107541  
GENERAL INFORMATION:  
APPLICANT: Arnold, Andrew  
TITLE OF INVENTION: Pradl Cyclin and its cdna  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
STREET: 1100 New York Avenue, N.W., Suite 600  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/460,744  
FILING DATE: 02-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: McConathy, Evelyn H.  
REGISTRATION NUMBER: 35,279  
REFERENCE/DOCKET NUMBER: 0609.4070005  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-2600  
TELEFAX: 202-371-2540  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 152 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-460-744-4

Query Match 59.3%; Score 32; DB 3; Length 152;  
Best Local Similarity 60.0%; Pred. No. 35;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 EEVVPXGMDY 10  
Db 20 EEVFLAMNY 29

RESULT 17  
US-07-667-711B-4  
Sequence 4, Application US/07667711B  
Patent No. 6110700  
GENERAL INFORMATION:  
APPLICANT: ARNOLD, ANDREW  
TITLE OF INVENTION: Pradl Cyclin and its cdna  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
STREET: 1100 NEW YORK AVE., NW, SUITE 600  
CITY: WASHINGTON  
STATE: DC  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30

;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/07/667,711B  
;; FILING DATE: 11-MAR-1991  
;; CLASSIFICATION: 435  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: MCPHAIL, DONALD R.  
;; REGISTRATION NUMBER: 35,811  
;; REFERENCE/DOCKET NUMBER: 0609.4070000  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (202) 371-2600  
;; TELEFAX: (202) 371-2540  
;; INFORMATION FOR SEQ ID NO: 4:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 152 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: not relevant  
;; MOLECULE TYPE: peptide  
US-07-667-711B-4

Query Match 59.3%; Score 32; DB 3; Length 152;  
Best Local Similarity 60.0%; Pred. No. 35;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMDY 10  
||| | :|  
Db 20 EEVFPAMNY 29

RESULT 18  
US-08-193-977-7  
; Sequence 7, Application US/08193977  
; Patent No. 5625031  
; GENERAL INFORMATION:  
; APPLICANT: WEBSTER, KEVIN R.  
; APPLICANT: COLEMAN, KEVIN G.  
; TITLE OF INVENTION: PEPTIDE INHIBITORS OF THE P33CDK2 AND  
; TITLE OF INVENTION: P34CDK2 CELL CYCLE REGULATORY KINASES AND HUMAN  
; TITLE OF INVENTION: PAPILLOMAVIRUS E7 ONCOPROTEIN  
; NUMBER OF SEQUENCES: 34  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: REED & ROBINS  
; STREET: 635 BRYANT STREET  
; CITY: PALO ALTO  
; STATE: CALIFORNIA  
; COUNTRY: UNITED STATES OF AMERICA  
; ZIP: 94301

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/193,977  
FILING DATE: 08-FEB-1994  
CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:  
NAME: ROBINS, ROBERTA L.  
REGISTRATION NUMBER: 33,208  
REFERENCE/DOCKET NUMBER: 5998-0016  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 617-8999  
TELEFAX: (415) 327-3231  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 173 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-193-977-7

Query Match 59.3%; Score 32; DB 1; Length 173;

Best Local Similarity 60.0%; Pred. No. 41;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
QY 1 EEVVPXGMDY 10  
||| | :|  
Db 55 EEVFPAMNY 64

RESULT 19  
US-08-464-517-21  
; Sequence 21, Application US/08464517  
; Patent No. 5869640  
; GENERAL INFORMATION:  
; APPLICANT: BEACH, David H.  
; TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO  
; NUMBER OF SEQUENCES: 50  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD  
; STREET: 60 State Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII(text)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/464,517  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/963,308  
FILING DATE: 16-OCT-1992  
APPLICATION NUMBER: US 07/888,178  
FILING DATE: 26-MAY-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/701,514  
FILING DATE: 16-MAY-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Matthew P. Vincent  
REGISTRATION NUMBER: 36,709  
REFERENCE/DOCKET NUMBER: MII-004C  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 227-7400  
TELEFAX: (617) 227-5941  
INFORMATION FOR SEQ ID NO: 21:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 189 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-464-517-21

Query Match 59.3%; Score 32; DB 2; Length 189;  
Best Local Similarity 60.0%; Pred. No. 46;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMDY 10  
||| | :|  
Db 74 EEVFPAMNY 83

RESULT 20  
US-08-246-361A-21  
; Sequence 21, Application US/08246361A  
; Patent No. 5998582  
; GENERAL INFORMATION:  
; APPLICANT: BEACH, David H.  
; TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO  
; NUMBER OF SEQUENCES: 50

;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: LAHIVE & COCKFIELD  
;; STREET: 60 State Street  
;; CITY: Boston  
;; STATE: MA  
;; COUNTRY: USA  
;; ZIP: 02109  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: ASCII(text)  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/246.361A  
;; FILING DATE: 19-MAY-1994  
;; CLASSIFICATION: 435  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/963,308  
;; FILING DATE: 16-OCT-1992  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/888,178  
;; FILING DATE: 26-MAY-1992  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/701,514  
;; FILING DATE: 16-MAY-1991  
;; NAME: Matthew P. Vincent  
;; REGISTRATION NUMBER: 36,709  
;; REFERENCE/DOCKET NUMBER: MII-004C  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (617) 227-7400  
;; TELEFAX: (617) 227-5941  
;; INFORMATION FOR SEQ ID NO: 21:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 189 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
US-08-246-361A-21

Query Match 59.3%; Score 32; DB 2; Length 189;  
Best Local Similarity 60.0%; Pred. No. 46;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMDY 10  
Db ||| | | |  
74 EEVFLPMNY 83

RESULT 21  
US-08-463-772-21  
; Sequence 21, Application US/08463772  
; GENERAL INFORMATION:  
; APPLICANT: BEACH, David H.  
; TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO  
; NUMBER OF SEQUENCES: 50  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD  
; STREET: 60 State Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII(text)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/463,772  
; FILING DATE:

;; CLASSIFICATION: 435  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/963,308  
;; FILING DATE: 16-OCT-1992  
;; APPLICATION NUMBER: US 07/888,178  
;; FILING DATE: 26-MAY-1992  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/701,514  
;; FILING DATE: 16-MAY-1991  
;; NAME: Matthew P. Vincent  
;; REGISTRATION NUMBER: 36,709  
;; REFERENCE/DOCKET NUMBER: MII-004C  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (617) 227-7400  
;; TELEFAX: (617) 227-5941  
;; INFORMATION FOR SEQ ID NO: 21:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 189 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
US-08-463-772-21

Query Match 59.3%; Score 32; DB 3; Length 189;  
Best Local Similarity 60.0%; Pred. No. 46;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMDY 10  
Db ||| | | |  
74 EEVFLPMNY 83

RESULT 22  
PCT-US93-05000-21  
; Sequence 21, Application PC/TUS9305000  
; GENERAL INFORMATION:  
; APPLICANT: MITOTIX  
; TITLE OF INVENTION: D-Type Cyclin and Uses Related Thereto  
; NUMBER OF SEQUENCES: 42  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
; STREET: Two Militia Drive  
; CITY: Lexington  
; STATE: Massachusetts  
; COUNTRY: US  
; ZIP: 02173  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US93/05000  
; FILING DATE: 19930525  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/07/888,178  
; FILING DATE: 26-MAY-1992  
; NAME: Granahan, Patricia  
; REGISTRATION NUMBER: 32,227  
; REFERENCE/DOCKET NUMBER: CSHL91-02A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-861-6240  
; TELEFAX: 616-861-9540  
; INFORMATION FOR SEQ ID NO: 21:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 189 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein



PCT-US93-05000-21

Query Match 59.3%; Score 32; DB 5; Length 189;  
Best Local Similarity 60.0%; Pred. No. 46;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMDY 10  
||| | | |  
Db 74 EEVFLPMNY 83

RESULT 23

US-08-926-842B-20  
; Sequence 20, Application US/08926842B  
; Patent No. 6030807  
; GENERAL INFORMATION:  
; APPLICANT: Sa-No. 6030807ueira, Isabel  
; APPLICANT: de Lencastre, Herminia  
; TITLE OF INVENTION: HIGHLY REGULABLE PROMOTER FOR HETEROLOGOUS GENE  
; TITLE OF INVENTION: EXPRESSION  
; NUMBER OF SEQUENCES: 64  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Klauber & Jackson  
; STREET: 411 Hackensack Avenue  
; CITY: Hackensack  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/926,842B  
; FILING DATE: 10-SEP-1997  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jackson Esq., David A.  
; REGISTRATION NUMBER: 26,742  
; REFERENCE/DOCKET NUMBER: 600-1-089 N  
; TELEPHONE: 201 487-5800  
; TELEFAX: 201 343-1684  
; TELEX: 133521  
; INFORMATION FOR SEQ ID NO: 20:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 231 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Escherichia coli  
US-08-926-842B-20

Query Match 59.3%; Score 32; DB 3; Length 231;  
Best Local Similarity 55.6%; Pred. No. 58;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 VVPXGMDYS 11  
: | | | | |  
Db 41 IKFSGVDYS 49

RESULT 24

US-08-464-517-22  
; Sequence 22, Application US/08464517  
; Patent No. 5869640  
; GENERAL INFORMATION:  
; APPLICANT: BEACH, David H.  
; TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO

; NUMBER OF SEQUENCES: 50  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD  
; STREET: 60 State Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII(text)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/464,517  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/963,308  
; FILING DATE: 16-OCT-1992  
; APPLICATION NUMBER: US 07/888,178  
; FILING DATE: 26-MAY-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/701,514  
; FILING DATE: 16-MAY-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Matthew P. Vincent  
; REGISTRATION NUMBER: 36,709  
; REFERENCE/DOCKET NUMBER: MII-004C  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 227-7400  
; TELEFAX: (617) 227-5941  
; INFORMATION FOR SEQ ID NO: 22:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 236 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-464-517-22

Query Match 59.3%; Score 32; DB 2; Length 236;  
Best Local Similarity 60.0%; Pred. No. 59;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMDY 10  
||| | | |  
Db 20 EEVFLPMNY 29

RESULT 25

US-08-246-361A-22  
; Sequence 22, Application US/08246361A  
; Patent No. 5998582  
; GENERAL INFORMATION:  
; APPLICANT: BEACH, David H.  
; TITLE OF INVENTION: D-TYPE CYCLIN AND USES RELATED THERETO  
; NUMBER OF SEQUENCES: 50  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD  
; STREET: 60 State Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII(text)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/246,361A  
; FILING DATE: 19-MAY-1994

```

; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/963,308
; FILING DATE: 16-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/888,178
; FILING DATE: 26-MAY-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/701,514
; FILING DATE: 16-MAY-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Matthew P. Vincent
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: MII-004C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 236 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-246-361A-22

Query Match          59.3%; Score 32; DB 2; Length 236;
Best Local Similarity 60.0%; Pred. No. 59;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY      1 EEVVPXGMDY 10
Db      20 EEVFPAMNY 29
```

Search completed: June 10, 2003, 13:51:34  
Job time : 10.6429 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 10, 2003, 13:46:50 ; Search time 15 Seconds  
(without alignments)  
75.710 Million cell updates/sec

Title: US-09-909-164-12  
Perfect score: 54  
Sequence: 1 EEVFXGMDYS 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 392085 seqs, 103240269 residues

Total number of hits satisfying chosen parameters: 392085

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 75 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	36	66.7	1049	9	US-10-217-700-10
2	36	66.7	1068	9	US-10-217-700-8
3	36	66.7	1081	9	US-10-217-700-4
4	36	66.7	1083	9	US-10-217-700-11
5	36	66.7	1084	9	US-10-217-700-9
6	35	64.8	440	9	US-09-813-408-27
7	34	63.0	1062	10	US-09-815-242-5111
8	34	63.0	3472	9	US-10-027-805-4
9	34	63.0	3472	9	US-10-034-623-4
10	34	63.0	3472	9	US-10-027-801-4
11	33	61.1	59	10	US-09-948-080-14
12	33	61.1	283	9	US-09-738-626-4881
13	33	61.1	299	10	US-09-815-242-10697
14	33	61.1	736	9	US-09-978-295A-526
15	33	61.1	736	9	US-09-978-697-526
16	33	61.1	736	9	US-09-978-192A-526
17	33	61.1	736	9	US-09-999-832A-526
18	33	61.1	736	9	US-09-978-189-526
19	33	61.1	736	9	US-10-174-590-420

20	33	61.1	736	9	US-10-176-758-420	Sequence 420, App
21	33	61.1	736	9	US-10-175-737-420	Sequence 420, App
22	33	61.1	736	9	US-10-173-706-420	Sequence 420, App
23	33	61.1	736	9	US-10-175-738-420	Sequence 420, App
24	33	61.1	736	9	US-10-175-752-420	Sequence 420, App
25	33	61.1	736	9	US-10-176-482-420	Sequence 420, App
26	33	61.1	736	9	US-10-176-757-420	Sequence 420, App
27	33	61.1	736	9	US-10-176-913-420	Sequence 420, App
28	33	61.1	736	9	US-10-180-552-420	Sequence 420, App
29	33	61.1	736	9	US-10-180-557-420	Sequence 420, App
30	33	61.1	736	9	US-10-173-700-420	Sequence 420, App
31	33	61.1	736	9	US-10-174-572-420	Sequence 420, App
32	33	61.1	736	9	US-10-174-579-420	Sequence 420, App
33	33	61.1	736	9	US-10-174-582-420	Sequence 420, App
34	33	61.1	736	9	US-10-174-588-420	Sequence 420, App
35	33	61.1	736	9	US-10-175-739-420	Sequence 420, App
36	33	61.1	736	9	US-10-175-740-420	Sequence 420, App
37	33	61.1	736	9	US-10-175-743-420	Sequence 420, App
38	33	61.1	736	9	US-10-176-488-420	Sequence 420, App
39	33	61.1	736	9	US-10-176-492-420	Sequence 420, App
40	33	61.1	736	9	US-10-176-747-420	Sequence 420, App
41	33	61.1	736	9	US-10-176-750-420	Sequence 420, App
42	33	61.1	736	9	US-10-176-985-420	Sequence 420, App
43	33	61.1	736	9	US-10-176-987-420	Sequence 420, App
44	33	61.1	736	9	US-10-176-991-420	Sequence 420, App
45	33	61.1	736	9	US-10-176-992-420	Sequence 420, App
46	33	61.1	736	9	US-10-176-993-420	Sequence 420, App
47	33	61.1	736	9	US-10-184-658-420	Sequence 420, App
48	33	61.1	736	9	US-10-173-695-420	Sequence 420, App
49	33	61.1	736	9	US-10-173-697-420	Sequence 420, App
50	33	61.1	736	9	US-10-173-705-420	Sequence 420, App
51	33	61.1	736	9	US-10-174-576-420	Sequence 420, App
52	33	61.1	736	9	US-10-174-585-420	Sequence 420, App
53	33	61.1	736	9	US-10-174-586-420	Sequence 420, App
54	33	61.1	736	9	US-10-175-747-420	Sequence 420, App
55	33	61.1	736	9	US-10-176-481-420	Sequence 420, App
56	33	61.1	736	9	US-10-176-485-420	Sequence 420, App
57	33	61.1	736	9	US-10-176-487-420	Sequence 420, App
58	33	61.1	736	9	US-10-176-493-420	Sequence 420, App
59	33	61.1	736	9	US-10-176-756-420	Sequence 420, App
60	33	61.1	736	9	US-10-176-919-420	Sequence 420, App
61	33	61.1	736	9	US-10-176-925-420	Sequence 420, App
62	33	61.1	736	9	US-10-176-978-420	Sequence 420, App
63	33	61.1	736	9	US-10-179-510-420	Sequence 420, App
64	33	61.1	736	9	US-10-180-543-420	Sequence 420, App
65	33	61.1	736	9	US-10-180-544-420	Sequence 420, App
66	33	61.1	736	9	US-10-180-546-420	Sequence 420, App
67	33	61.1	736	9	US-10-180-547-420	Sequence 420, App
68	33	61.1	736	9	US-10-180-549-420	Sequence 420, App
69	33	61.1	736	9	US-10-180-555-420	Sequence 420, App
70	33	61.1	736	9	US-10-180-559-420	Sequence 420, App
71	33	61.1	736	9	US-10-181-000-420	Sequence 420, App
72	33	61.1	736	9	US-10-181-010-420	Sequence 420, App
73	33	61.1	736	9	US-10-183-012-420	Sequence 420, App
74	33	61.1	736	9	US-10-183-010-420	Sequence 420, App
75	33	61.1	736	9	US-10-184-614-420	Sequence 420, App

ALIGNMENTS

RESULT 1  
US-10-217-700-10  
; Sequence 10, Application US10217700  
; Publication No. US20030070191A1  
; GENERAL INFORMATION:  
; APPLICANT: Haigler, Candace H.  
; APPLICANT: Holaday, A. Scott  
; TITLE OF INVENTION: TRANSGENIC FIBER PRODUCING PLANTS WITH INCREASED  
; ; TITLE OF INVENTION: EXPRESSION OF SUCROSE PHOSPHATE SYNTHASE  
; ; FILE REFERENCE: 201304/1000  
; ; CURRENT APPLICATION NUMBER: US/10/217,700  
; ; CURRENT FILING DATE: 2002-08-12

; EARLIER APPLICATION NUMBER: 09/394,272  
; EARLIER FILING DATE: 1999-09-10  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 10  
; LENGTH: 1049  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
US-10-217-700-10

Query Match  
Best Local Similarity 66.7%; Score 36; DB 9; Length 1049;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMDYS 11  
|:| |||:  
Db 436 VIPPGMDFS 444

## RESULT 2

US-10-217-700-8  
; Sequence 8, Application US/10217700  
; Publication No. US20030070191A1  
; GENERAL INFORMATION:  
; APPLICANT: Haigler, Candace H.  
; TITLE OF INVENTION: TRANSGENIC FIBER PRODUCING PLANTS WITH INCREASED  
; FILE REFERENCE: 201304/1000  
; CURRENT APPLICATION NUMBER: US/10/217,700  
; EARLIER FILING DATE: 2002-08-12  
; EARLIER APPLICATION NUMBER: 09/394,272  
; EARLIER FILING DATE: 1999-09-10  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 8  
; LENGTH: 1068  
; TYPE: PRT  
; ORGANISM: Zea mays  
US-10-217-700-8

Query Match  
Best Local Similarity 66.7%; Score 36; DB 9; Length 1068;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMDYS 11  
|:| |||:  
Db 435 VIPPGMDFS 443

## RESULT 3

US-10-217-700-4  
; Sequence 4, Application US/10217700  
; Publication No. US20030070191A1  
; GENERAL INFORMATION:  
; APPLICANT: Haigler, Candace H.  
; TITLE OF INVENTION: TRANSGENIC FIBER PRODUCING PLANTS WITH INCREASED  
; FILE REFERENCE: 201304/1000  
; CURRENT APPLICATION NUMBER: US/10/217,700  
; EARLIER FILING DATE: 2002-08-12  
; EARLIER APPLICATION NUMBER: 09/394,272  
; EARLIER FILING DATE: 1999-09-10  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 4  
; LENGTH: 1081  
; TYPE: PRT  
; ORGANISM: Craterostigma plantagineum  
US-10-217-700-4

Query Match  
Best Local Similarity 66.7%; Score 36; DB 9; Length 1081;

Best Local Similarity 66.7%; Pred. No. 1e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMDYS 11  
|:| |||:  
Db 445 VIPPGMDFS 453

## RESULT 4

US-10-217-700-11  
; Sequence 11, Application US/10217700  
; Publication No. US20030070191A1  
; GENERAL INFORMATION:  
; APPLICANT: Haigler, Candace H.  
; TITLE OF INVENTION: TRANSGENIC FIBER PRODUCING PLANTS WITH INCREASED  
; FILE REFERENCE: 201304/1000  
; CURRENT APPLICATION NUMBER: US/10/217,700  
; EARLIER FILING DATE: 2002-08-12  
; EARLIER APPLICATION NUMBER: 09/394,272  
; EARLIER FILING DATE: 1999-09-10  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 11  
; LENGTH: 1083  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
US-10-217-700-11

Query Match  
Best Local Similarity 66.7%; Score 36; DB 9; Length 1083;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMDYS 11  
|:| |||:  
Db 483 VIPPGMDFS 491

## RESULT 5

US-10-217-700-9  
; Sequence 9, Application US/10217700  
; Publication No. US20030070191A1  
; GENERAL INFORMATION:  
; APPLICANT: Haigler, Candace H.  
; TITLE OF INVENTION: TRANSGENIC FIBER PRODUCING PLANTS WITH INCREASED  
; FILE REFERENCE: 201304/1000  
; CURRENT APPLICATION NUMBER: US/10/217,700  
; EARLIER FILING DATE: 2002-08-12  
; EARLIER APPLICATION NUMBER: 09/394,272  
; EARLIER FILING DATE: 1999-09-10  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 9  
; LENGTH: 1084  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
US-10-217-700-9

Query Match  
Best Local Similarity 66.7%; Score 36; DB 9; Length 1084;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VVPXGMDYS 11  
|:| |||:  
Db 453 VIPPGMDFS 461

## RESULT 6

US-09-813-408-27  
; Sequence 27, Application US/09813408

Publication No. US20030049619A1  
 GENERAL INFORMATION:  
 APPLICANT: Delagrave, Simon  
 APPLICANT: Marrs, Barry  
 TITLE OF INVENTION: Methods For The Synthesis Of Polynucleotides And Combinatorial Libraries  
 TITLE OF INVENTION: Of Polynucleotides  
 FILE REFERENCE: HER0041  
 CURRENT APPLICATION NUMBER: US/09/813,408  
 CURRENT FILING DATE: 2001-03-21  
 NUMBER OF SEQ ID NOS: 85  
 SOFTWARE: PatentIn version 3.0  
 SEQ ID NO 27  
 LENGTH: 440  
 TYPE: PRT  
 ORGANISM: Aeropyrum pernix  
 US-09-813-408-27

Query Match 64.8%; Score 35; DB 9; Length 440;  
 Best Local Similarity 66.7%; Pred. No. 58;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 EVVPXGMDY 10  
 |||||  
 Db 120 EVLPWGVY 128

## RESULT 7

US-09-815-242-5111  
 Sequence 5111, Application US/09815242  
 Patent No. US20020061569A1  
 GENERAL INFORMATION:  
 APPLICANT: Haselbeck, Robert  
 APPLICANT: Ohlsen, Karl L.  
 APPLICANT: Zyskind, Judith W.  
 APPLICANT: Wall, Daniel  
 APPLICANT: Trawick, John D.  
 APPLICANT: Carr, Grant J.  
 APPLICANT: Yamamoto, Robert T.  
 APPLICANT: Xu, H. Howard

TITLE OF INVENTION: Identification of Essential Genes in  
 TITLE OF INVENTION: Prokaryotes

FILE REFERENCE: ELITRA.011A

CURRENT APPLICATION NUMBER: US/09/815,242

CURRENT FILING DATE: 2001-03-21

PRIOR APPLICATION NUMBER: 60/191,078

PRIOR FILING DATE: 2000-03-21

PRIOR APPLICATION NUMBER: 60/206,848

PRIOR FILING DATE: 2000-05-23

PRIOR APPLICATION NUMBER: 60/207,727

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: 60/242,578

PRIOR FILING DATE: 2000-10-23

PRIOR APPLICATION NUMBER: 60/253,625

PRIOR FILING DATE: 2000-11-27

PRIOR APPLICATION NUMBER: 60/257,931

PRIOR FILING DATE: 2000-12-22

PRIOR APPLICATION NUMBER: 60/269,308

PRIOR FILING DATE: 2001-02-16

NUMBER OF SEQ ID NOS: 14110

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 5111

LENGTH: 1062

TYPE: PRT

ORGANISM: Pseudomonas aeruginosa

US-09-815-242-5111

Query Match 63.0%; Score 34; DB 10; Length 1062;  
 Best Local Similarity 85.7%; Pred. No. 2.4e+02;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 PXGMDYS 11  
 |||||  
 Db 321 PQGMDYS 327

## RESULT 8

US-10-027-806-4  
 Sequence 4, Application US/10027806  
 Patent No. US20020160476A1

GENERAL INFORMATION:

APPLICANT: Swanson, Ronald V.

APPLICANT: Feldman, Robert A.

APPLICANT: Schleper, Christa

TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM CENARCHAEUM SYMBIOSUM

FILE REFERENCE: DCRP.002A

CURRENT APPLICATION NUMBER: US/10/027,806

CURRENT FILING DATE: 2001-12-21

PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/408,020

PRIOR FILING DATE: EARLIER FILING DATE: 1999-09-29

NUMBER OF SEQ ID NOS: 123

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 4

LENGTH: 3472

TYPE: PRT

ORGANISM: Cenarchaeum symbiosum

US-10-027-806-4

Query Match 63.0%; Score 34; DB 9; Length 3472;

Best Local Similarity 45.5%; Pred. No. 9.1e+02;

Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMDYS 11

|||||

Db 2294 EDVIPRGISFS 2304

## RESULT 9

US-10-034-623-4

Sequence 4, Application US/10034623

Publication No. US20020198365A1

GENERAL INFORMATION:

APPLICANT: Swanson, Ronald V.

APPLICANT: Feldman, Robert A.

APPLICANT: Schleper, Christa

TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM CENARCHAEUM SYMBIOSUM

FILE REFERENCE: DCRP.002A

CURRENT APPLICATION NUMBER: US/10/034,623

CURRENT FILING DATE: 2001-12-21

PRIOR APPLICATION NUMBER: 09/408,020

PRIOR FILING DATE: 1999-09-29

PRIOR APPLICATION NUMBER: 60/102,294

PRIOR FILING DATE: 1998-09-29

NUMBER OF SEQ ID NOS: 123

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 4

LENGTH: 3472

TYPE: PRT

ORGANISM: Cenarchaeum symbiosum

US-10-034-623-4

Query Match 63.0%; Score 34; DB 9; Length 3472;

Best Local Similarity 45.5%; Pred. No. 9.1e+02;

Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMDYS 11

|||||

Db 2294 EDVIPRGISFS 2304

## RESULT 10

US-10-027-801-4

Sequence 4, Application US/10027801

Publication No. US20030054364A1

GENERAL INFORMATION:

APPLICANT: Swanson, Ronald V.

APPLICANT: Feldman, Robert A.

; APPLICANT: Schleper, Christa  
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS FROM CENARCHAEUM SYMBIOSIUM  
; FILE REFERENCE: DCOBP.002A  
; CURRENT APPLICATION NUMBER: US/10/027,801  
; PRIOR FILING DATE: 2001-12-21  
; PRIOR APPLICATION NUMBER: EARLIER FILING DATE: 1999-09-29  
; PRIOR FILING DATE: 2001-12-21  
; NUMBER OF SEQ ID NOS: 123  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 4  
; LENGTH: 3472  
; TYPE: PRT  
; ORGANISM: Cenarchaeum symbiosum  
US-10-027-801-4

Query Match 63.0%; Score 34; DB 9; Length 3472;  
Best Local Similarity 45.5%; Pred. No. 9.1e+02;  
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMDYS 11  
Db 2294 EDVIPRGISFS 2304  
|:|:| |:|

## RESULT 11

US-09-948-080-14  
; Sequence 14, Application US/09948080  
; Patent No. US20020102702A1  
; GENERAL INFORMATION:  
; APPLICANT: VAN DER OSTEN, CLAUS  
; APPLICANT: HALKIER, TORDEN  
; APPLICANT: ANDERSEN, CARSTEN  
; APPLICANT: BAUDITZ, PETER  
; APPLICANT: HANSEN, PETER KAMP  
; TITLE OF INVENTION: PROTEASE VARIANTS AND COMPOSITIONS  
; FILE REFERENCE: 4946,200-US  
; CURRENT APPLICATION NUMBER: US/09/948,080  
; CURRENT FILING DATE: 2001-09-06  
; PRIOR APPLICATION NUMBER: US/08/963,851  
; PRIOR FILING DATE: 1997-11-04  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 14  
; LENGTH: 59  
; TYPE: PRT  
; ORGANISM: Enterococcus faecalis  
US-09-948-080-14

Query Match 61.1%; Score 33; DB 10; Length 59;  
Best Local Similarity 45.5%; Pred. No. 15;  
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMDYS 11  
Db 38 EXHIPGLEY 48  
|:|:| |:|

## RESULT 12

US-09-738-626-4881  
; Sequence 4881, Application US/09738626  
; Publication No. US20020197605A1  
; GENERAL INFORMATION:  
; APPLICANT: NAKAGAWA, SATOSHI  
; APPLICANT: MIZOGUCHI, HIROSHI  
; APPLICANT: ANDO, SEIKO  
; APPLICANT: HAYASHI, MIKIO  
; APPLICANT: OCHIAI, KEIKO  
; APPLICANT: YOKOI, HARUHIKO  
; APPLICANT: TATEISHI, NAKO  
; APPLICANT: SENOH, AKIHIRO  
; APPLICANT: IKEDA, MASATO  
; APPLICANT: OKAKI, AKIO  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES

; FILE REFERENCE: 249-125  
; CURRENT APPLICATION NUMBER: US/09/738,626  
; CURRENT FILING DATE: 2000-12-18  
; PRIOR APPLICATION NUMBER: JP 99/377484  
; PRIOR FILING DATE: 1999-12-16  
; PRIOR APPLICATION NUMBER: JP 00/159162  
; PRIOR FILING DATE: 2000-04-07  
; PRIOR APPLICATION NUMBER: JP 00/280988  
; PRIOR FILING DATE: 2000-08-03  
; NUMBER OF SEQ ID NOS: 7059  
; SOFTWARE: PatentIn ver. 3.0  
; SEQ ID NO 4881  
; LENGTH: 283  
; TYPE: PRT  
; ORGANISM: Corynebacterium glutamicum  
US-09-738-626-4881

Query Match 61.1%; Score 33; DB 9; Length 283;  
Best Local Similarity 75.0%; Pred. No. 88;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 VPXGMDYS 11  
Db 56 VPAGADYS 63  
|:|:| |:|

## RESULT 13

US-09-815-242-10697  
; Sequence 10697, Application US/09815242  
; Patent No. US20020061569A1  
; GENERAL INFORMATION:  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari L.  
; APPLICANT: Zyskind, Judith W.  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John D.  
; APPLICANT: Carr, Grant J.  
; APPLICANT: Yamamoto, Robert T.  
; APPLICANT: Xu, H. Howard  
; TITLE OF INVENTION: Identification of Essential Genes in  
; FILE REFERENCE: ELITRA.011A  
; CURRENT APPLICATION NUMBER: US/09/815,242  
; CURRENT FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; NUMBER OF SEQ ID NOS: 14110  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 10697  
; LENGTH: 299  
; TYPE: PRT  
; ORGANISM: Enterococcus faecalis  
US-09-815-242-10697

Query Match 61.1%; Score 33; DB 10; Length 299;  
Best Local Similarity 40.0%; Pred. No. 94;  
Matches 4; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 EEVVPXGMDY 10  
Db 218 EQITPTGIEY 227  
|:|:| |:|

## RESULT 14

US-09-978-295A-526

Sequence 526, Application US/09978295A

Patent No. US20020156006A1

GENERAL INFORMATION:

APPLICANT: Ashkenazi, Avi  
APPLICANT: Baker Kevin P.  
APPLICANT: Botstein, David  
APPLICANT: Desnoyers, Luc  
APPLICANT: Eaton, Dan  
APPLICANT: Ferrara, Napoleon  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Fong, Sherman  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Gerber, Hanspeter  
APPLICANT: Gerritsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Grimaldi, J. Christopher  
APPLICANT: Gurney, Austin L.  
APPLICANT: Hillan, Kenneth J.  
APPLICANT: Kijavlin, Ivar J.  
APPLICANT: Kuo, Sophia S.  
APPLICANT: Napier, Mary A.  
APPLICANT: Pan, James;  
APPLICANT: Paoni, Nicholas F.  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Shelton, David L.  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William I.  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
FILE REFERENCE: P2630P1C11  
CURRENT APPLICATION NUMBER: US/09/978,295A  
CURRENT FILING DATE: 2001-10-15  
PRIOR APPLICATION NUMBER: 09/918585  
PRIOR FILING DATE: 2001-07-30  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/064249  
PRIOR FILING DATE: 1997-11-03  
PRIOR APPLICATION NUMBER: 60/065311  
PRIOR FILING DATE: 1997-11-13  
PRIOR APPLICATION NUMBER: 60/066364  
PRIOR FILING DATE: 1997-11-21  
PRIOR APPLICATION NUMBER: 60/077450  
PRIOR FILING DATE: 1998-03-10  
PRIOR APPLICATION NUMBER: 60/077632  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077641  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077649  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077791  
PRIOR FILING DATE: 1998-03-12  
PRIOR APPLICATION NUMBER: 60/078004  
PRIOR FILING DATE: 1998-03-13  
PRIOR APPLICATION NUMBER: 60/078886  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078936  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078939  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/079294  
PRIOR FILING DATE: 1998-03-25  
PRIOR APPLICATION NUMBER: 60/079656  
PRIOR FILING DATE: 1998-03-26

PRIOR APPLICATION NUMBER: 60/079664  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079689  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079663  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079728  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079786  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079920  
PRIOR FILING DATE: 1998-03-30  
PRIOR APPLICATION NUMBER: 60/079923  
PRIOR FILING DATE: 1998-03-30  
PRIOR APPLICATION NUMBER: 60/080105  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080107  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080165  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080194  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080327  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080328  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080333  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/081049  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081071  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081195  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081203  
PRIOR FILING DATE: 1998-04-09  
PRIOR APPLICATION NUMBER: 60/081229  
PRIOR FILING DATE: 1998-04-09  
PRIOR APPLICATION NUMBER: 60/081955  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081817  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081819  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081952  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081838  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/082568  
PRIOR FILING DATE: 1998-04-21  
PRIOR APPLICATION NUMBER: 60/082569  
PRIOR FILING DATE: 1998-04-21  
PRIOR APPLICATION NUMBER: 60/082704  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082804  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082700  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082797  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082796  
PRIOR FILING DATE: 1998-04-23  
PRIOR APPLICATION NUMBER: 60/083336  
PRIOR FILING DATE: 1998-04-27  
PRIOR APPLICATION NUMBER: 60/083322  
PRIOR FILING DATE: 1998-04-28  
PRIOR APPLICATION NUMBER: 60/083392  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083495

```

; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083496
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083499
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083545
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083554
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083558
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083559
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083500
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083742
; PRIOR FILING DATE: 1998-04-30
; PRIOR APPLICATION NUMBER: 60/084366
; PRIOR FILING DATE: 1998-05-05
; PRIOR APPLICATION NUMBER: 60/084414
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084441
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084637
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084639
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084640
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084598
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084627
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084643
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/085339
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085338
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085323
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085582
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085700
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085689
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085580
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085573
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085704
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697

```

Query Match 61.1%; Score 33; DB 9; Length 736;  
 Best Local Similarity 70.0%; Pred. No. 2.6e+02;

Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

QY 1 EEVVPXGMDY 10
Db 331 EPVVVYGM DY 340

```

## RESULT 15

US-09-978-697-526  
 ; Sequence 526, Application US/09978697  
 ; Patent No. US20020169284A1.  
 ; GENERAL INFORMATION:

```

; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James;
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Shelton, David L.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2630PIC27
; CURRENT APPLICATION NUMBER: US/09/978,697
; CURRENT FILING DATE: 2001-10-16
; PRIOR APPLICATION NUMBER: 09/918585
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/064249
; PRIOR FILING DATE: 1997-11-03
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066364
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/077450
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: 60/077632
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077641
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077649
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077791
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/078004
; PRIOR FILING DATE: 1998-03-13
; PRIOR APPLICATION NUMBER: 60/078886
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078936
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078939
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079664
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079663
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079728

```



;; PRIOR FILING DATE: 1998-03-27  
;; PRIOR APPLICATION NUMBER: 60/079786  
;; PRIOR FILING DATE: 1998-03-27  
;; PRIOR APPLICATION NUMBER: 60/079920  
;; PRIOR FILING DATE: 1998-03-30  
;; PRIOR APPLICATION NUMBER: 60/079923  
;; PRIOR FILING DATE: 1998-03-30  
;; PRIOR APPLICATION NUMBER: 60/080105  
;; PRIOR FILING DATE: 1998-03-31  
;; PRIOR APPLICATION NUMBER: 60/080107  
;; PRIOR FILING DATE: 1998-03-31  
;; PRIOR APPLICATION NUMBER: 60/080165  
;; PRIOR FILING DATE: 1998-03-31  
;; PRIOR APPLICATION NUMBER: 60/080194  
;; PRIOR FILING DATE: 1998-03-31  
;; PRIOR APPLICATION NUMBER: 60/080327  
;; PRIOR FILING DATE: 1998-04-01  
;; PRIOR APPLICATION NUMBER: 60/080328  
;; PRIOR FILING DATE: 1998-04-01  
;; PRIOR APPLICATION NUMBER: 60/080333  
;; PRIOR FILING DATE: 1998-04-01  
;; PRIOR APPLICATION NUMBER: 60/080334  
;; PRIOR FILING DATE: 1998-04-01  
;; PRIOR APPLICATION NUMBER: 60/081070  
;; PRIOR FILING DATE: 1998-04-08  
;; PRIOR APPLICATION NUMBER: 60/081049  
;; PRIOR FILING DATE: 1998-04-08  
;; PRIOR APPLICATION NUMBER: 60/081071  
;; PRIOR FILING DATE: 1998-04-08  
;; PRIOR APPLICATION NUMBER: 60/081195  
;; PRIOR FILING DATE: 1998-04-08  
;; PRIOR APPLICATION NUMBER: 60/081203  
;; PRIOR FILING DATE: 1998-04-09  
;; PRIOR APPLICATION NUMBER: 60/081229  
;; PRIOR FILING DATE: 1998-04-09  
;; PRIOR APPLICATION NUMBER: 60/081955  
;; PRIOR FILING DATE: 1998-04-15  
;; PRIOR APPLICATION NUMBER: 60/081817  
;; PRIOR FILING DATE: 1998-04-15  
;; PRIOR APPLICATION NUMBER: 60/081819  
;; PRIOR FILING DATE: 1998-04-15  
;; PRIOR APPLICATION NUMBER: 60/081952  
;; PRIOR FILING DATE: 1998-04-15  
;; PRIOR APPLICATION NUMBER: 60/081838  
;; PRIOR FILING DATE: 1998-04-15  
;; PRIOR APPLICATION NUMBER: 60/082568  
;; PRIOR FILING DATE: 1998-04-21  
;; PRIOR APPLICATION NUMBER: 60/082569  
;; PRIOR FILING DATE: 1998-04-21  
;; PRIOR APPLICATION NUMBER: 60/082704  
;; PRIOR FILING DATE: 1998-04-22  
;; PRIOR APPLICATION NUMBER: 60/082804  
;; PRIOR FILING DATE: 1998-04-22  
;; PRIOR APPLICATION NUMBER: 60/082700  
;; PRIOR FILING DATE: 1998-04-22  
;; PRIOR APPLICATION NUMBER: 60/082797  
;; PRIOR FILING DATE: 1998-04-22  
;; PRIOR APPLICATION NUMBER: 60/082796  
;; PRIOR FILING DATE: 1998-04-23  
;; PRIOR APPLICATION NUMBER: 60/083336  
;; PRIOR FILING DATE: 1998-04-27  
;; PRIOR APPLICATION NUMBER: 60/083322  
;; PRIOR FILING DATE: 1998-04-28  
;; PRIOR APPLICATION NUMBER: 60/083392  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/083495  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/083496  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/083499  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/083545  
;; PRIOR FILING DATE: 1998-04-29

;; PRIOR APPLICATION NUMBER: 60/083554  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/083558  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/083559  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/083500  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/083742  
;; PRIOR FILING DATE: 1998-04-30  
;; PRIOR APPLICATION NUMBER: 60/084366  
;; PRIOR FILING DATE: 1998-05-05  
;; PRIOR APPLICATION NUMBER: 60/084414  
;; PRIOR FILING DATE: 1998-05-06  
;; PRIOR APPLICATION NUMBER: 60/084441  
;; PRIOR FILING DATE: 1998-05-06  
;; PRIOR APPLICATION NUMBER: 60/084637  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/084639  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/084640  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/084598  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/084600  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/084627  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/084643  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/085339  
;; PRIOR FILING DATE: 1998-05-13  
;; PRIOR APPLICATION NUMBER: 60/085338  
;; PRIOR FILING DATE: 1998-05-13  
;; PRIOR APPLICATION NUMBER: 60/085323  
;; PRIOR FILING DATE: 1998-05-13  
;; PRIOR APPLICATION NUMBER: 60/085582  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085700  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085689  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085579  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085580  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085573  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085704  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085697

Query Match 61.1%; Score 33; DB 9; Length 736;  
Best Local Similarity 70.0%; Pred. No. 2.6e+02;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMDY 10  
Db 331 EPVVVYGM DY 340

RESULT 16  
US-09-978-192A-526  
; Sequence 526, Application US/09978192A  
; Patent No. US20020177553A1  
; GENERAL INFORMATION:  
; APPLICANT: Ashkenazi, Avi  
; APPLICANT: Baker Kevin P.  
; APPLICANT: Botstein, David  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Eaton, Dan  
; APPLICANT: Ferrara, Napoleon  
; APPLICANT: Filvaroff, Ellen

APPLICANT: Fong, Sherman  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Gerber, Hanspeter  
APPLICANT: Gerritsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Grimaldi, J. Christopher  
APPLICANT: Gurney, Austin L.  
APPLICANT: Hillan, Kenneth J.  
APPLICANT: Kijavlin, Ivar J.  
APPLICANT: Kuo, Sophia S.  
APPLICANT: Napier, Mary A.  
APPLICANT: Pan, James  
APPLICANT: Peoni, Nicholas F.  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Shelton, David L.  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William I.  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
FILE REFERENCE: P2630P1C9  
CURRENT APPLICATION NUMBER: US/09/978,192A  
CURRENT FILING DATE: 2001-10-15  
PRIOR APPLICATION NUMBER: 09/918585  
PRIOR FILING DATE: 2001-07-30  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/064249  
PRIOR FILING DATE: 1997-11-03  
PRIOR APPLICATION NUMBER: 60/065311  
PRIOR FILING DATE: 1997-11-13  
PRIOR APPLICATION NUMBER: 60/066364  
PRIOR FILING DATE: 1997-11-21  
PRIOR APPLICATION NUMBER: 60/077450  
PRIOR FILING DATE: 1998-03-10  
PRIOR APPLICATION NUMBER: 60/077632  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077641  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077649  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077791  
PRIOR FILING DATE: 1998-03-12  
PRIOR APPLICATION NUMBER: 60/078004  
PRIOR FILING DATE: 1998-03-13  
PRIOR APPLICATION NUMBER: 60/078886  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078936  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078939  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/079294  
PRIOR FILING DATE: 1998-03-25  
PRIOR APPLICATION NUMBER: 60/079656  
PRIOR FILING DATE: 1998-03-26  
PRIOR APPLICATION NUMBER: 60/079664  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079689  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079663  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079728  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079786  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079920  
PRIOR FILING DATE: 1998-03-30  
PRIOR APPLICATION NUMBER: 60/079923  
PRIOR FILING DATE: 1998-03-30  
PRIOR APPLICATION NUMBER: 60/080105  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080107  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080165  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080194  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080327  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080328  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080333  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080334  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/081070  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081049  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081071  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081195  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081203  
PRIOR FILING DATE: 1998-04-09  
PRIOR APPLICATION NUMBER: 60/081229  
PRIOR FILING DATE: 1998-04-09  
PRIOR APPLICATION NUMBER: 60/081955  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081817  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081819  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081952  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081838  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/082568  
PRIOR FILING DATE: 1998-04-21  
PRIOR APPLICATION NUMBER: 60/082569  
PRIOR FILING DATE: 1998-04-21  
PRIOR APPLICATION NUMBER: 60/082704  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082804  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082700  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082797  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082796  
PRIOR FILING DATE: 1998-04-23  
PRIOR APPLICATION NUMBER: 60/083336  
PRIOR FILING DATE: 1998-04-27  
PRIOR APPLICATION NUMBER: 60/083322  
PRIOR FILING DATE: 1998-04-28  
PRIOR APPLICATION NUMBER: 60/083392  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083495  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083496  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083499  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083545  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083554  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083558  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083559  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083500

;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/083742  
;; PRIOR FILING DATE: 1998-04-30  
;; PRIOR APPLICATION NUMBER: 60/084366  
;; PRIOR FILING DATE: 1998-05-05  
;; PRIOR APPLICATION NUMBER: 60/084414  
;; PRIOR FILING DATE: 1998-05-06  
;; PRIOR APPLICATION NUMBER: 60/084441  
;; PRIOR FILING DATE: 1998-05-06  
;; PRIOR APPLICATION NUMBER: 60/084637  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/084639  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/084640  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/084598  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/084600  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/084627  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/084643  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/085339  
;; PRIOR FILING DATE: 1998-05-13  
;; PRIOR APPLICATION NUMBER: 60/085338  
;; PRIOR FILING DATE: 1998-05-13  
;; PRIOR APPLICATION NUMBER: 60/085323  
;; PRIOR FILING DATE: 1998-05-13  
;; PRIOR APPLICATION NUMBER: 60/085582  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085700  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085689  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085579  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085580  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085573  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085704  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085697

Query Match 61.1%; Score 33; DB 9; Length 736;  
Best Local Similarity 70.0%; Pred. No. 2.6e+02;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMDY 10  
Db 331 EPVVVYGM DY 340

RESULT 17  
US-09-999-832A-526  
; Sequence 526, Application US/0999832A  
; Publication No. US20020192706A1  
; GENERAL INFORMATION:

;; APPLICANT: Ashkenazi, Avi  
;; APPLICANT: Baker Kevin P.  
;; APPLICANT: Botstein, David  
;; APPLICANT: Desnoyers, Luc  
;; APPLICANT: Eaton, Dan  
;; APPLICANT: Ferrara, Napoleon  
;; APPLICANT: Filvaroff, Ellen  
;; APPLICANT: Fong, Sherman  
;; APPLICANT: Gao, Wei-Qiang  
;; APPLICANT: Gerber, Hanspeter  
;; APPLICANT: Gerritsen, Mary E.  
;; APPLICANT: Goddard, Audrey  
;; APPLICANT: Godowski, Paul J.  
;; APPLICANT: Grimaldi, J. Christopher

;; APPLICANT: Gurney, Austin L.  
;; APPLICANT: Hillan, Kenneth J.  
;; APPLICANT: Kljavin, Ivar J.  
;; APPLICANT: Kuo, Sophia S.  
;; APPLICANT: Napier, Mary A.  
;; APPLICANT: Pan, James;  
;; APPLICANT: Paoni, Nicholas F.  
;; APPLICANT: Roy, Margaret Ann  
;; APPLICANT: Shelton, David L.  
;; APPLICANT: Stewart, Timothy A.  
;; APPLICANT: Tumas, Daniel  
;; APPLICANT: Williams, P. Mickey  
;; APPLICANT: Wood, William I.  
;; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
;; FILE REFERENCE: P2630PIC63  
;; CURRENT FILING DATE: 2001-10-24  
;; PRIOR APPLICATION NUMBER: US/09/999,832A  
;; PRIOR FILING DATE: 2001-07-30  
;; PRIOR APPLICATION NUMBER: 60/062250  
;; PRIOR FILING DATE: 1997-10-17  
;; PRIOR APPLICATION NUMBER: 60/064249  
;; PRIOR FILING DATE: 1997-11-03  
;; PRIOR APPLICATION NUMBER: 60/065311  
;; PRIOR FILING DATE: 1997-11-13  
;; PRIOR APPLICATION NUMBER: 60/066364  
;; PRIOR FILING DATE: 1997-11-21  
;; PRIOR APPLICATION NUMBER: 60/077450  
;; PRIOR FILING DATE: 1998-03-10  
;; PRIOR APPLICATION NUMBER: 60/077632  
;; PRIOR FILING DATE: 1998-03-11  
;; PRIOR APPLICATION NUMBER: 60/077641  
;; PRIOR FILING DATE: 1998-03-11  
;; PRIOR APPLICATION NUMBER: 60/077649  
;; PRIOR FILING DATE: 1998-03-11  
;; PRIOR APPLICATION NUMBER: 60/077791  
;; PRIOR FILING DATE: 1998-03-12  
;; PRIOR APPLICATION NUMBER: 60/078004  
;; PRIOR FILING DATE: 1998-03-13  
;; PRIOR APPLICATION NUMBER: 60/078886  
;; PRIOR FILING DATE: 1998-03-20  
;; PRIOR APPLICATION NUMBER: 60/078936  
;; PRIOR FILING DATE: 1998-03-20  
;; PRIOR APPLICATION NUMBER: 60/078910  
;; PRIOR FILING DATE: 1998-03-20  
;; PRIOR APPLICATION NUMBER: 60/078939  
;; PRIOR FILING DATE: 1998-03-20  
;; PRIOR APPLICATION NUMBER: 60/079294  
;; PRIOR FILING DATE: 1998-03-25  
;; PRIOR APPLICATION NUMBER: 60/079656  
;; PRIOR FILING DATE: 1998-03-26  
;; PRIOR APPLICATION NUMBER: 60/079664  
;; PRIOR FILING DATE: 1998-03-27  
;; PRIOR APPLICATION NUMBER: 60/079689  
;; PRIOR FILING DATE: 1998-03-27  
;; PRIOR APPLICATION NUMBER: 60/079663  
;; PRIOR FILING DATE: 1998-03-27  
;; PRIOR APPLICATION NUMBER: 60/079728  
;; PRIOR FILING DATE: 1998-03-27  
;; PRIOR APPLICATION NUMBER: 60/079786  
;; PRIOR FILING DATE: 1998-03-27  
;; PRIOR APPLICATION NUMBER: 60/079920  
;; PRIOR FILING DATE: 1998-03-30  
;; PRIOR APPLICATION NUMBER: 60/079923  
;; PRIOR FILING DATE: 1998-03-30  
;; PRIOR APPLICATION NUMBER: 60/080105  
;; PRIOR FILING DATE: 1998-03-31  
;; PRIOR APPLICATION NUMBER: 60/080107  
;; PRIOR FILING DATE: 1998-03-31  
;; PRIOR APPLICATION NUMBER: 60/080165  
;; PRIOR FILING DATE: 1998-03-31  
;; PRIOR APPLICATION NUMBER: 60/080194

; PRIOR FILING DATE: 1998-03-31  
 ; PRIOR APPLICATION NUMBER: 60/080327  
 ; PRIOR FILING DATE: 1998-04-01  
 ; PRIOR APPLICATION NUMBER: 60/080328  
 ; PRIOR FILING DATE: 1998-04-01  
 ; PRIOR APPLICATION NUMBER: 60/080333  
 ; PRIOR FILING DATE: 1998-04-01  
 ; PRIOR APPLICATION NUMBER: 60/080334  
 ; PRIOR FILING DATE: 1998-04-01  
 ; PRIOR APPLICATION NUMBER: 60/081070  
 ; PRIOR FILING DATE: 1998-04-08  
 ; PRIOR APPLICATION NUMBER: 60/081049  
 ; PRIOR FILING DATE: 1998-04-08  
 ; PRIOR APPLICATION NUMBER: 60/081071  
 ; PRIOR FILING DATE: 1998-04-08  
 ; PRIOR APPLICATION NUMBER: 60/081195  
 ; PRIOR FILING DATE: 1998-04-08  
 ; PRIOR APPLICATION NUMBER: 60/081203  
 ; PRIOR FILING DATE: 1998-04-09  
 ; PRIOR APPLICATION NUMBER: 60/081229  
 ; PRIOR FILING DATE: 1998-04-09  
 ; PRIOR APPLICATION NUMBER: 60/081955  
 ; PRIOR FILING DATE: 1998-04-15  
 ; PRIOR APPLICATION NUMBER: 60/081817  
 ; PRIOR FILING DATE: 1998-04-15  
 ; PRIOR APPLICATION NUMBER: 60/081819  
 ; PRIOR FILING DATE: 1998-04-15  
 ; PRIOR APPLICATION NUMBER: 60/081952  
 ; PRIOR FILING DATE: 1998-04-15  
 ; PRIOR APPLICATION NUMBER: 60/081838  
 ; PRIOR FILING DATE: 1998-04-15  
 ; PRIOR APPLICATION NUMBER: 60/082568  
 ; PRIOR FILING DATE: 1998-04-21  
 ; PRIOR APPLICATION NUMBER: 60/082569  
 ; PRIOR FILING DATE: 1998-04-21  
 ; PRIOR APPLICATION NUMBER: 60/082704  
 ; PRIOR FILING DATE: 1998-04-22  
 ; PRIOR APPLICATION NUMBER: 60/082804  
 ; PRIOR FILING DATE: 1998-04-22  
 ; PRIOR APPLICATION NUMBER: 60/082700  
 ; PRIOR FILING DATE: 1998-04-22  
 ; PRIOR APPLICATION NUMBER: 60/082797  
 ; PRIOR FILING DATE: 1998-04-22  
 ; PRIOR APPLICATION NUMBER: 60/082796  
 ; PRIOR FILING DATE: 1998-04-23  
 ; PRIOR APPLICATION NUMBER: 60/083336  
 ; PRIOR FILING DATE: 1998-04-27  
 ; PRIOR APPLICATION NUMBER: 60/083322  
 ; PRIOR FILING DATE: 1998-04-28  
 ; PRIOR APPLICATION NUMBER: 60/083392  
 ; PRIOR FILING DATE: 1998-04-29  
 ; PRIOR APPLICATION NUMBER: 60/083495  
 ; PRIOR FILING DATE: 1998-04-29  
 ; PRIOR APPLICATION NUMBER: 60/083496  
 ; PRIOR FILING DATE: 1998-04-29  
 ; PRIOR APPLICATION NUMBER: 60/083499  
 ; PRIOR FILING DATE: 1998-04-29  
 ; PRIOR APPLICATION NUMBER: 60/083545  
 ; PRIOR FILING DATE: 1998-04-29  
 ; PRIOR APPLICATION NUMBER: 60/083554  
 ; PRIOR FILING DATE: 1998-04-29  
 ; PRIOR APPLICATION NUMBER: 60/083558  
 ; PRIOR FILING DATE: 1998-04-29  
 ; PRIOR APPLICATION NUMBER: 60/083559  
 ; PRIOR FILING DATE: 1998-04-29  
 ; PRIOR APPLICATION NUMBER: 60/083500  
 ; PRIOR FILING DATE: 1998-04-29  
 ; PRIOR APPLICATION NUMBER: 60/083742  
 ; PRIOR FILING DATE: 1998-04-30  
 ; PRIOR APPLICATION NUMBER: 60/084366  
 ; PRIOR FILING DATE: 1998-05-05  
 ; PRIOR APPLICATION NUMBER: 60/084414  
 ; PRIOR FILING DATE: 1998-05-06

; PRIOR APPLICATION NUMBER: 60/084441  
 ; PRIOR FILING DATE: 1998-05-06  
 ; PRIOR APPLICATION NUMBER: 60/084637  
 ; PRIOR FILING DATE: 1998-05-07  
 ; PRIOR APPLICATION NUMBER: 60/084639  
 ; PRIOR FILING DATE: 1998-05-07  
 ; PRIOR APPLICATION NUMBER: 60/084640  
 ; PRIOR FILING DATE: 1998-05-07  
 ; PRIOR APPLICATION NUMBER: 60/084598  
 ; PRIOR FILING DATE: 1998-05-07  
 ; PRIOR APPLICATION NUMBER: 60/084600  
 ; PRIOR FILING DATE: 1998-05-07  
 ; PRIOR APPLICATION NUMBER: 60/084627  
 ; PRIOR FILING DATE: 1998-05-07  
 ; PRIOR APPLICATION NUMBER: 60/084643  
 ; PRIOR FILING DATE: 1998-05-07  
 ; PRIOR APPLICATION NUMBER: 60/085339  
 ; PRIOR FILING DATE: 1998-05-13  
 ; PRIOR APPLICATION NUMBER: 60/085338  
 ; PRIOR FILING DATE: 1998-05-13  
 ; PRIOR APPLICATION NUMBER: 60/085323  
 ; PRIOR FILING DATE: 1998-05-13  
 ; PRIOR APPLICATION NUMBER: 60/085582  
 ; PRIOR FILING DATE: 1998-05-15  
 ; PRIOR APPLICATION NUMBER: 60/085700  
 ; PRIOR FILING DATE: 1998-05-15  
 ; PRIOR APPLICATION NUMBER: 60/085689  
 ; PRIOR FILING DATE: 1998-05-15  
 ; PRIOR APPLICATION NUMBER: 60/085579  
 ; PRIOR FILING DATE: 1998-05-15  
 ; PRIOR APPLICATION NUMBER: 60/085580  
 ; PRIOR FILING DATE: 1998-05-15  
 ; PRIOR APPLICATION NUMBER: 60/085573  
 ; PRIOR FILING DATE: 1998-05-15  
 ; PRIOR APPLICATION NUMBER: 60/085704  
 ; PRIOR FILING DATE: 1998-05-15  
 ; PRIOR APPLICATION NUMBER: 60/085697

Query Match 61.1%; Score 33; DB 9; Length 736;  
 Best Local Similarity 70.0%; Pred. No. 2.6e+02;  
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 BEVVPXGMDY 10  
 | | | | |  
 Db 331 EPVVVYGM DY 340

RESULT 18  
 US-09-978-189-526  
 ; Sequence 526, Application US/09978189  
 ; Publication No. US20030004102A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Ashkenazi, Avi  
 ; APPLICANT: Baker Kevin P.  
 ; APPLICANT: Botstein, David  
 ; APPLICANT: Desnoyers, Luc  
 ; APPLICANT: Eaton, Dan  
 ; APPLICANT: Ferrara, Napoleon  
 ; APPLICANT: Filvaroff, Ellen  
 ; APPLICANT: Fong, Sherman  
 ; APPLICANT: Gao, Wei-Qiang  
 ; APPLICANT: Gerber, Hanspeter  
 ; APPLICANT: Gerritsen, Mary E.  
 ; APPLICANT: Goddard, Audrey  
 ; APPLICANT: Godowski, Paul J.  
 ; APPLICANT: Grimaldi, J. Christopher  
 ; APPLICANT: Gurney, Austin L.  
 ; APPLICANT: Hillan, Kenneth J.  
 ; APPLICANT: Kljavin, Ivar J.  
 ; APPLICANT: Kuo, Sophia S.  
 ; APPLICANT: Napier, Mary A.  
 ; APPLICANT: Pan, James;  
 ; APPLICANT: Paoni, Nicholas F.

APPLICANT: Roy, Margaret Ann  
APPLICANT: Shelton, David L.  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William I.  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
TITLE OF INVENTION: Acids Encoding the Same  
FILE REFERENCE: P2630PIC7  
CURRENT APPLICATION NUMBER: US/09/978,189  
CURRENT FILING DATE: 2001-10-15  
PRIOR APPLICATION NUMBER: 09/918585  
PRIOR FILING DATE: 2001-07-30  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/064249  
PRIOR FILING DATE: 1997-11-03  
PRIOR APPLICATION NUMBER: 60/065311  
PRIOR FILING DATE: 1997-11-13  
PRIOR APPLICATION NUMBER: 60/066364  
PRIOR FILING DATE: 1997-11-21  
PRIOR APPLICATION NUMBER: 60/077450  
PRIOR FILING DATE: 1998-03-10  
PRIOR APPLICATION NUMBER: 60/077632  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077641  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077649  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077791  
PRIOR FILING DATE: 1998-03-12  
PRIOR APPLICATION NUMBER: 60/078004  
PRIOR FILING DATE: 1998-03-13  
PRIOR APPLICATION NUMBER: 60/078886  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078936  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078939  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/079294  
PRIOR FILING DATE: 1998-03-25  
PRIOR APPLICATION NUMBER: 60/079656  
PRIOR FILING DATE: 1998-03-26  
PRIOR APPLICATION NUMBER: 60/079664  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079689  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079663  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079728  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079786  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079920  
PRIOR FILING DATE: 1998-03-30  
PRIOR APPLICATION NUMBER: 60/079923  
PRIOR FILING DATE: 1998-03-30  
PRIOR APPLICATION NUMBER: 60/080105  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080107  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080165  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080194  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080327  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080328  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080333  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080334  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/081070  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081049  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081071  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081195  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081203  
PRIOR FILING DATE: 1998-04-09  
PRIOR APPLICATION NUMBER: 60/081229  
PRIOR FILING DATE: 1998-04-09  
PRIOR APPLICATION NUMBER: 60/081955  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081817  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081819  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081952  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081838  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/082568  
PRIOR FILING DATE: 1998-04-21  
PRIOR APPLICATION NUMBER: 60/082569  
PRIOR FILING DATE: 1998-04-21  
PRIOR APPLICATION NUMBER: 60/082704  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082804  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082700  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082797  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082796  
PRIOR FILING DATE: 1998-04-23  
PRIOR APPLICATION NUMBER: 60/083336  
PRIOR FILING DATE: 1998-04-27  
PRIOR APPLICATION NUMBER: 60/083322  
PRIOR FILING DATE: 1998-04-28  
PRIOR APPLICATION NUMBER: 60/083392  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083495  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083496  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083499  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083545  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083554  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083558  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083559  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083500  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083742  
PRIOR FILING DATE: 1998-04-30  
PRIOR APPLICATION NUMBER: 60/084366  
PRIOR FILING DATE: 1998-05-05  
PRIOR APPLICATION NUMBER: 60/084414  
PRIOR FILING DATE: 1998-05-06  
PRIOR APPLICATION NUMBER: 60/084441  
PRIOR FILING DATE: 1998-05-06  
PRIOR APPLICATION NUMBER: 60/084637  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084639  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084640

;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/084598  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/084600  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/084627  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/084643  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/085339  
;; PRIOR FILING DATE: 1998-05-13  
;; PRIOR APPLICATION NUMBER: 60/085338  
;; PRIOR FILING DATE: 1998-05-13  
;; PRIOR APPLICATION NUMBER: 60/085323  
;; PRIOR FILING DATE: 1998-05-13  
;; PRIOR APPLICATION NUMBER: 60/085582  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085700  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085689  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085579  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085580  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085573  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085704  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085697

Query Match 61.1%; Score 33; DB 9; Length 736;  
Best Local Similarity 70.0%; Pred. No. 2.6e+02;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMDY 10  
DB 331 EPVVVYGM DY 340

RESULT 19  
US-10-174-590-420  
; Sequence 420, Application US/10174590  
; Publication No. US20030008352A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Chen, Jian  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Pan, James  
; APPLICANT: Smith, Victoria  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Wood, William I.  
; APPLICANT: Zhang, Zemin

;; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
;; FILE REFERENCE: P3430RIC42  
;; CURRENT APPLICATION NUMBER: US/10/174,590  
;; CURRENT FILING DATE: 2002-06-18  
;; Prior application removed - See File Wrapper or Palm  
;; NUMBER OF SEQ ID NOS: 612  
;; SEQ ID NO 420  
;; LENGTH: 736  
;; TYPE: PRT  
;; ORGANISM: Homo Sapien  
US-10-174-590-420

Query Match 61.1%; Score 33; DB 9; Length 736;  
Best Local Similarity 70.0%; Pred. No. 2.6e+02;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMDY 10  
DB 331 EPVVVYGM DY 340

## RESULT 20

US-10-176-758-420  
; Sequence 420, Application US/10176758  
; Publication No. US20030008353A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Chen, Jian  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Pan, James  
; APPLICANT: Smith, Victoria  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Wood, William I.  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3430RIC104  
; CURRENT APPLICATION NUMBER: US/10/176,758  
; CURRENT FILING DATE: 2002-06-21  
; Prior application removed - See File Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 612  
; SEQ ID NO 420  
; LENGTH: 736  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-176-758-420

Query Match 61.1%; Score 33; DB 9; Length 736;  
Best Local Similarity 70.0%; Pred. No. 2.6e+02;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 EEVVPXGMDY 10  
DB 331 EPVVVYGM DY 340

## RESULT 21

US-10-175-737-420  
; Sequence 420, Application US/10175737  
; Publication No. US20030013153A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Chen, Jian  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Pan, James  
; APPLICANT: Smith, Victoria  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Wood, William I.  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3430RIC50  
; CURRENT APPLICATION NUMBER: US/10/175,737  
; CURRENT FILING DATE: 2002-06-19  
; Prior application removed - See File Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 612  
; SEQ ID NO 420  
; LENGTH: 736  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-175-737-420

Query Match 61.1%; Score 33; DB 9; Length 736;